Título original: A SURVEY OF MODERN ALGEBRA

Traducido de la 12.ª edición inglesa (1953), con autorización de

THE MACMILLAN COMPANY

Primera edición en español, 1954

1.ª Reimpresión, 1960
2.ª Reimpresión, 1963

Depósito Legal, B. 3419 -- 1960

PRINTED IN SPAIN
IMPRESO EN ESPAÑA

Editado por EDITORIAL VICENS — VIVES, Avda. de Sarriá, 136, Barcelona, 17

Impreso por GRAFICAS INSTAR, San Gervasio de Cassolas, 79, Barcelona, 6

Prólogo del traductor

En la literatura científica en castellano carecíase de un tratado didáctico de Álgebra moderna. Creemos, sinceramente, que la traducción que ofrecemos será muy útil al público estudioso de nuestra patria y de los países de habla castellana, entre otras razones, por la importancia de la materia tratada, por la eficiencia y valor de la obra y, asimismo, por la personalidad de sus autores, los prestigiosos profesores G. Birkhoff y Saunders MacLane.

Existe, como es bien sabido, una marcada distinción entre el Álgebra clásica y la llamada Álgebra moderna. Ésta ha surgido como superación de aquélla para atender a cuestiones originadas en los más diversos campos de las Matemáticas. Al Álgebra moderna corresponde, en efecto, lograr una fundamentación y sistematización precisas para el estudio de los conjuntos en que se definen operaciones (la naturaleza de los elementos del conjunto y de las operaciones entre ellos es indiferente; lo esencial son las reglas de cálculo). Se comprende, desde luego, que al corresponder al Álgebra moderna un problema tan general como el señalado, sus métodos y resultados penetren en los más variados dominios de la Matemática y aun de la Ciencia en general. Efectivamente, ésta tiende, cada vez más, a la abstracción conceptual, y, por ende, a utilizar los recursos y el lenguaje de aquélla.

La exposición de una doctrina tan esencialmente abstracta sin que resulte árida para el lector no especialmente preparado, ha de ser, necesariamente, obra maestra de método y claridad didáctica. Esto es lo que han conseguido los autores del presente tratado. Buena prueba de ello son las doce sucesivas ediciones del texto americano, así como el universal consentimiento que han hallado, en el mundo entero, entre profesores universitarios, técnicos industriales y estudiosos interesados por la ciencia contemporánea.

El secreto de semejante éxito estriba, sin duda, en la eminente capacidad científica y la extrema habilidad didáctica de los autores, de acrisolada autoridad mundial en el campo de la investigación. G. Birkhoff y Saunders MacLane, profesores de Matemáticas de las Universidades de Harvard y Chicago, ocupan, en efecto, uno de los primeros rangos en ambos aspectos en la constelación científica mundial.

Nos consideraremos altamente satisfechos si con la versión de esta obra al castellano logramos una nueva etapa de progreso en conocimiento tan esencial para el desarrollo de la ciencia y la técnica modernas.

R. Rodríguez Vidal

Prefacio

La más notable característica del Algebra Moderna es la investigación de las propiedades teóricas de sistemas formales dados, tales como grupos, anillos, campos y espacios vectoriales. Al escribir el presente texto nos hemos esforzado en destacar este método «abstracto», pero dejándonos guiar por una interpretación muy amplia del significado del Algebra Moderna. Nos parece que gran parte de este significado está en la interpretación personal del tema. Por consiguiente, hemos intentado en todo caso expresar el apoyo conceptual de las varias definiciones. Esto se ha logrado ilustrando cada nueva idea con numerosos ejemplos, lo más familiares posible. Ello parece de especial importancia en un texto elemental, porque sirve para destacar el hecho de que todos los conceptos abstractos nacen del análisis de situaciones concretas.

Para desarrollar en el estudiante la facultad de pensar por sí mismo mediante estos nuevos conceptos, hemos incluído gran variedad de ejercicios sobre cada tema. Algunos de estos ejercicios son puramente de control, pero algunos exploran más profundamente en los nuevos conceptos y otros ofrecen desarrollos teóricos adicionales. Los ejercicios del último tipo ofrecen el importante servicio de familiarizar al estudiante con la construcción de demostraciones formales. La selección de ejercicios permitirá al profesor adaptar el texto a estudiantes de diversos grados de madurez.

El Álgebra Moderna, además, facilita una reinterpretación de los resultados del Álgebra Clásica, dándoles mayor generalidad y unidad. Por lo tanto, en vez de omitir estos resultados, hemos procurado incorporarlos de modo sistemático al marco de ideas del Álgebra Moderna.

También hemos tenido en cuenta el hecho de que para muchos estudiantes, el valor del Álgebra está en sus aplicaciones a otras

ciencias: Análisis Superior, Geometría, Física y Filosofía. Esto ha influído en la atención que hemos prestado a los campos real y complejo, a los grupos de transformaciones contrastados con los grupos abstractos, a las matrices simétricas y reducidas a forma diagonal, a la clasificación afín y métrica de las cuádricas y, finalmente, en la inclusión del álgebra de clases, teoría de redes (lattices) y números transfinitos, cuestiones que son de importancia en la Lógica matemática y en la moderna Teoría de Funciones.

La cuestión de las matrices y sus aplicaciones es de gran importancia; hemos procurado darla con completa independencia, sin acudir a otras materias que puedan obscurecerla. Es deseable que no se tengan sólo presentes las operaciones formales con matrices, sino también su interpretación (muchas veces descuidada) como transformaciones lineales. Hemos mantenido siempre a la vista esta interpretación, destacando las oportunas propiedades en el espacio vectorial. Las propiedades de las matrices regulares se desarrollan por el método directo, sin el auxilio de los determinantes, pero el profesor que prefiera este último método, puede anticipar el estudio completo de las propiedades básicas de los determinantes, que se expone en el Capítulo X.

El libro comienza con dos capítulos dedicados a los números enteros y racionales, desarrollando los correspondientes conceptos abstractos de dominio de integridad y campo (o cuerpo), así como algo de la teoría elemental de números. Los tres capítulos siguientes tratan, en lenguaje moderno, los puntos básicos de la teoría de ecuaciones, incluyendo la estructura de los sistemas de números reales y complejos. En el Capítulo VI se introduce, con multitud de ejemplos, el concepto fundamental de grupo. Éste se aplica, en los Capítulos VII a X, a los espacios vectoriales y matrices. Los Capítulos XI y XII se dedican a los conjuntos y Aritmética transfinita. Finalmente, los tres últimos capítulos proporcionan una introducción al Álgebra conmutativa y Aritmética: ampliaciones de un campo, números algebraicos, ideales y grupos de Galois. La teoría de Galois es desarrollada, sin apelar a las funciones simétricas, por el elegante método de Artin.

Muchos de estos capítulos son mutuamente independientes; por ejemplo, el capítulo sobre teoría de grupos puede darse inmediatamente después del Capítulo I, mientras que los relativos a ideales y campos (al principio de los Capítulos XIII y XIV) pue-

den ser estudiados a continuación del capítulo sobre espacios vectoriales.

Esperamos que estas ordenaciones puedan hacer el libro adaptable a cursos de diverso tipo. Un año escolar típico puede darse con los Capítulos I a X (omitiendo algunas secciones, que se han señalado con un asterisco), más algunos puntos escogidos de los otros capítulos. Tal curso supondría sólo conocimientos normales de Álgebra clásica. Los estudiantes que dominen la teoría de las ecuaciones pueden suprimir, o repasar ligeramente, los Capítulos III-V, dedicando, en cambio, más atención a los últimos del libro.

Un curso más breve, destinado a proporcionar los conocimientos que se utilizan en la Física, puede basarse en los Capítulos VI a X, dándole siempre a la palabra «campo» un significado restringido al campo real o al complejo. Un curso breve de Álgebra abstracta puede hacerse con los Capítulos I-IV, VI, VII, VIII, XI, XIII y XIV. Muchas otras distribuciones son posibles.

Nuestra deuda con varios textos magistrales de Álgebra Moderna es evidente ; hemos también consultado con provecho los apuntes del Prof. Nathan Jacobson. Nuestra gratitud al Dr. Walther Leighton, Prof. Clifford Bell y Prof. A. D. Campbell, que utilizaron una primera redacción de este libro y nos ofrecieron sus comentarios y sugestiones. Apreciamos la ayuda de Mrs. Saunders MacLane, que contribuyó con su labor de secretaria ; de Theodore Singer, que comprobó los Ejercicios, y de otros muchos estudiantes y colegas que nos han ayudado en la preparación de este libro.

GARRETT BIRKHOFF
SAUNDERS MACLANE

Cambridge.

INDICE

CAPÍTULO I

Los enteros

1. Introducción

El Álgebra Moderna ha puesto de manifiesto toda la variedad e importancia de los sistemas matemáticos posibles. En esta obra construiremos y estudiaremos varios de tales sistemas; pero el más importante de todos ellos es también el más antiguo, y consiste en el conjunto de los números enteros positivos (números naturales). Otro sistema, algo más amplio, es el formado por todos los números enteros (positivos, negativos y cero). Comenzaremos con la discusión de este último, porque presenta estrechas semejanzas con otros sistemas más generales introducidos por el Álgebra Moderna.

En vez de intentar definir directamente lo que son los números enteros, comenzaremos por suponer que estos enteros, cualquier cosa que ellos sean, vienen obligados a satisfacer ciertas leyes algebraicas fundamentales. Estas leyes o hipótesis se eligen de tal modo que sea posible demostrar, a partir de ellas, todas las otras propiedades de los enteros. Se dirá que un sistema tal de hipótesis constituye los *axiomas* o *postulados* fundamentales del sistema de los enteros. Veremos más adelante (Cap. I, § 13) que cualquier sistema de entes que los satisfaga es, salvo la notación, equivalente a dicho sistema de enteros.

Algunos de los postulados que definen a los enteros pueden ser reemplazados por otros igualmente convenientes, tales como la posibilidad de la división, o la posibilidad de resolver la ecuación $x^2 = -1$. La construcción de nuevos sistemas que tengan estas propiedades nos lleva inevitablemente a considerar nuevas clases de números, como las fracciones o los números reales (esto es, todos los que tienen desarrollo decimal indefinido) o los números complejos (en los que interviene $i = \sqrt{-1}$).

El método de los postulados (o método axiomático) es común-

mente empleado en la enseñanza elemental de la Geometría plana y del espacio. Pero mientras en la primera, por ejemplo, se desarrollan las consecuencias de un conjunto adecuado de postulados, que se suponen aplicados a los puntos de un plano, y sólo a ellos, el desarrollo sistemático del Álgebra implica un sistema de postulados que son aplicables simultáneamente a muchos sistemas algebraicos. Por ejemplo, la ley distributiva $a(b+c)=ab+ac$, y varias de las otras reglas familiares del cálculo algebraico, se aplican igualmente a las diversas familias de los números enteros, fraccionarios, reales o complejos. Por este motivo es conveniente disponer de nombres genéricos que sean aplicables a todos los sistemas que cumplen un conjunto dado de postulados. Este es el origen de varias locuciones, que serán definidas más adelante, tales como «dominio de integridad», «anillo», «grupo», «espacio vectorial», etc., todas de importancia primordial en el Álgebra Moderna.

2. Dominios de integridad

Analicemos en primer lugar las propiedades de la adición y multiplicación en el sistema de los enteros ordinarios: 0, ± 1, ± 2, ± 3, ... Las siguientes leyes quedan satisfechas por todos los enteros a, b, c:

	Adición	*Multiplicación*
Ley conmutativa	$a+b=b+a$	$ab=ba$
Ley asociativa	$a+(b+c)=(a+b)+c$	$a(bc)=(ab)c$
Ley distributiva	$a(b+c)=ab+ac$	

Las leyes conmutativa y asociativa resultan tan conocidas que se las utiliza sin mención expresa. Por ejemplo, con $a+b+c$ se designan indistintamente los números $a+(b+c)$ y $(a+b)+c$, sin hacer distinción del orden con que se comienza a sumar.

El número cero tiene la propiedad característica de que deja inalterado cualquier otro con el que se suma. Diremos que el cero es un «elemento idéntico» para la adición. Por evidente analogía formal, el 1 será llamado elemento idéntico para la multiplicación. Brevemente:

Identidades: $a+0=a \quad a \cdot 1=a$ para todo, a

El opuesto $-a$ de un entero a tiene la propiedad de ser

Aditivo inverso (opuesto): $a+(-a)=0$

Esta ley es equivalente a afirmar que la ecuación $a+x=0$ tiene como solución $x=-a$. Finalmente, los elementos no cero que aparezcan como factores en los dos miembros de una igualdad, pueden ser suprimidos o reducidos, es decir:

Ley de simplificación: Si $c \neq 0$ y $ca=cb$, es $a=b$

Estas mismas leyes algebraicas se aplican igualmente a otras clases de números; son válidas, por ejemplo, para los racionales y para los reales. Incluso se aplican, no sólo a los números, sino también a los polinomios. Por lo tanto, será conveniente llamar «dominio de integridad» a todo conjunto de elementos que satisfagan estas leyes. He aquí el enunciado completo de este convenio:

DEFINICIÓN. *Un dominio de integridad D es cualquier conjunto de elementos entre los cuales están definidas dos operaciones, llamadas adición y multiplicación, con las siguientes propiedades:*

a) *Cada par de elementos* a *y* b *de D determinan unívocamente una suma* a+b *y un producto* a·b *en D, de modo que sean válidas la ley distributiva, las dos leyes asociativas y las dos conmutativas;*

b) *D contiene dos elementos distintos, 0 (cero) y 1 (uno), que son idénticos para la adición y la multiplicación respectivamente;*

c) *Para cada* a *en D, la ecuación* a+x=0 *tiene en D una solución,* x=—a.

d) *Es válida la ley de simplificación de productos.*

El conjunto J de todos los enteros, el conjunto R de todos los números racionales y el conjunto R* de todos los números reales son ejemplos de dominios de integridad. Otro ejemplo, menos corriente, es el de todos los números de la forma $a+b\sqrt{3}$, donde a y b son enteros ordinarios. La suma y el producto de dos números de esta forma puede ser siempre puesto en la misma forma, pues

$$(a+b\sqrt{3})+(c+d\sqrt{3})=(a+c)+(b+d)\sqrt{3}$$

$$(a+b\sqrt{3})\,(c+d\sqrt{3})=(ac+3bd)+(ad+bc)\sqrt{3}$$

Es fácil comprobar que con estas fórmulas se cumplen las leyes distributiva, asociativa y conmutativa; se observa también que $0=0+0\sqrt{3}$ es el cero, y $1=1+0\sqrt{3}$ es la unidad. Además, la ecuación $(a+b\sqrt{3})+x=0$, tiene por solución $x=-a-b\sqrt{3}$, que es un número de la forma que estamos considerando. Por último, la ley de simplificación de productos también es válida, como se demostrará al final de esta sección.

Dados estos varios ejemplos de dominios de integridad, cualquier resultado válido en uno de ellos, que hayamos deducido mediante sólo las hipótesis o postulados de la definición, será también válido en todos los otros. Esta observación es aplicable a las propiedades usuales de la adición y la multiplicación, que ahora vamos a deducir a partir de los precedentes postulados.

Regla 1. Hay tan sólo un elemento cero (elemento idéntico en la adición). En efecto, si hubiese un segundo $0'$, con $a+0'=a$ y $b+0=b$, y haciendo $a=0$, $b=0'$, y aplicando la ley conmutativa, sería

$$0'=0'+0=0+0'=0 \qquad\qquad 0'=0$$

como decíamos. Se prueba de un modo análogo que hay un solo elemento unidad (idéntico en la multiplicación).

Regla 2. Es válida la ley de simplificación para la suma; esto es,

$$a+b=a+c \qquad \text{implica} \qquad b=c$$

En efecto: sumando $-a$ en ambos miembros de la igualdad dada y aplicando la ley asociativa, es:

$$b=(-a+a)+b=-a+(a+b)=-a+(a+c)=(-a+a)+c=0+c=c$$

(c. q. d.)

Regla 3. La sustracción es posible y unívoca, lo cual quiere decir que la ecuación $a+x=b$ tiene siempre para cada a y b en D una solución única. En efecto, por sustitución directa, se comprueba que $(-a)+b$ es una solución; además, si hubiese dos soluciones x e y, entonces $a+x=b$, $a+y=b$, y la ley de simplificación de la suma daría $y=x$. La solución única $(-a)+b$, se representa comúnmente por $b-a$.

Regla 4. El elemento $-a$ resulta caracterizado como la única solución de la ecuación $a+x=0$.

Regla 5. El elemento idéntico en la adición tiene la propiedad $a \cdot 0 = 0$ para cualquier a. En efecto, por la ley distributiva,

$$a \cdot a + a \cdot 0 = a(a+0) = a \cdot a = a \cdot a + 0$$

y cancelando $a \cdot a$ en ambos miembros, resulta lo enunciado.

Regla 6. Dos elementos negativos dan producto positivo: $(-a) \cdot (-b) = ab$. Para establecerlo consideremos la triple suma $ab + a(-b) + (-a) \cdot (-b)$. Con los dos primeros sumandos, por la ley distributiva, resulta: $a[b + (-b)] = a \cdot 0 = 0$; luego toda la expresión equivale a $(-a) \cdot (-b)$. Análogamente, los dos últimos términos suman $[a + (-a)](-b) = 0 \cdot (-b) = 0$, y la expresión es igual a ab. Por consiguiente, $(-a)(-b) = ab$, c. q. d.

Caso particular es la chocante fórmula $(-1)(-1) = 1$.

Regla 7. Existe una *ley asociativa general*, según la cual los términos de una suma o los factores de un producto pueden ser agrupados arbitrariamente. Para probarlo, se aplica varias veces sucesivas la ley asociativa ordinaria, y se obtiene que la suma de n números dados es independiente del modo en que se les agrupa entre paréntesis (*). Por esta razón, una suma así se escribe de ordinario sin ningún paréntesis: $a_1 + \ldots + a_n$. Lo mismo se dice para el producto.

Regla 8. La ley conmutativa hace posible alterar el orden de los términos; hay, pues, una ley asociativa y conmutativa general, según la cual, la suma (o el producto) de n términos dados tiene el mismo valor, cualquiera que sea el modo de ordenarlos y de agruparlos.

Otras muchas leyes sencillas se deducen del anterior sistema de postulados, y algunas se proponen a continuación como ejercicio.

Otra ley algebraica fundamental es la que se utiliza en la resolución de ecuaciones cuadráticas, cuando se afirma que $(x+2)(x-3) = 0$, significa o que $x+2=0$ o que $x-3=0$. La ley general se formula así:

(1) si $a \cdot b = 0$, o es $a = 0$ o es $b = 0$

La demostración es inmediata, por la ley de simplificación. Supongamos, en efecto, que el primer factor a no es cero. Entonces $a \cdot b = 0 = a \cdot 0$, y a puede suprimirse; luego $b = 0$. Inversamente, la ley de simplificación se deduce de esta propiedad y de los res-

(*) Esta demostración envuelve el empleo de la inducción completa. (Cfr. § 5.)

tantes postulados, pues si $a \neq 0$, $a \cdot b = a \cdot c$ significa que $ab - ac = a(b-c) = 0$, y de aquí, por esta última propiedad, $b - c = 0$, y $b = c$. Por lo tanto, resulta :

TEOREMA 1. *La ley de simplificación para el producto es lógicamente equivalente a afirmar que el producto de dos factores no nulos es distinto de cero.*

Los elementos no nulos a y b, cuyo producto ab es igual a cero, son llamados a veces «divisores del cero» o «divisores nulos» ; así que la ley de simplificación de productos en un dominio de integridad D equivale a afirmar que D no contiene divisores nulos.

Como aplicación de este teorema, demostraremos la ley de simplificación en el dominio considerado antes, constituído por todos los números $a + b\sqrt{3}$ (a y b enteros), admitiendo que la ecuación $x^2 = 3y^2$ no puede tener como solución dos enteros x e y no nulos (lo que demostraremos en Cap. III). De aquí resulta que $a + b\sqrt{3}$ es cero, sólo si $a = b = 0$, ya que $a + b\sqrt{3} = 0$ equivale a $a = -b\sqrt{3}$, o sea, $a^2 = 3b^2$.

Supongamos ahora que en los números considerados hubiese divisores de cero, como

$$(a + b\sqrt{3})(c + d\sqrt{3}) = (ac + 3bd) + (ad + bc)\sqrt{3} = 0$$

Por la observación anterior sería $ac + 3bd = 0$, $ad + bc = 0$; si multiplicamos la primera por d, la segunda por c y restamos, resultará $b(3d^2 - c^2) = 0$. Como la ley de simplificación vale para los enteros, o es $b = 0$ o es $3d^2 - c^2 = 0$. Si $3d^2 - c^2 = 0$, debe ser $d = c = 0$, por lo dicho. Si $b = 0$, las dos ecuaciones anteriores dan $ac = ad = 0$, así que o es $a = 0$ o es $c = d = 0$. En cualquier caso, uno de los supuestos divisores de cero, $a + b\sqrt{3}$ o $c + d\sqrt{3}$, resulta ser también igual a cero.

EJERCICIOS

1. Demostrar que en todo dominio de integridad son válidas las siguientes reglas:

a) $(-a) = (-1)a$ b) $-(a+b) = (-a) + (-b)$

c) $(-a)b = a(-b) = -(ab)$ d) $-(-a) = a$

e) $a(b-c) = ab - ac$ f) $(a-b) + (b-c) = a - c$

g) Existe un solo elemento 1 tal que $a \cdot 1 = a$ para todo a.

h) Los únicos elementos «idempotentes», esto es, que satisfacen a la igualdad $xx = x$, son 0 y 1.

2. Demostrar la regla 6 a partir del Ejercicio 1 a) y del caso particular $(-1)\cdot(-1)=1$.
3. Probar que las siguientes reglas valen en todo dominio de integridad:
   a) $(a-b)+(c-d)=(a+c)-(b+d)$
   b) $(a-b)-(c-d)=(a+d)-(b+c)$
   c) $(a-b)(c-d)=(ac+bd)-(ad+bc)$
   d) $(a-b)=(c-d)$ si, y sólo si, $a+d=b+c$
4. ¿Cuáles de los siguientes conjuntos de números son dominio de integridad?
   a) Todos los enteros pares.
   b) Todos los enteros impares.
   c) Todos los números de la forma $a+b\sqrt{2}$, con a y b enteros.
   d) Todos los números reales de la forma $a+b\cdot 5^{1/4}$, donde a y b son enteros.
   e) Todos los números reales de la forma $a+b\cdot 9^{1/4}$, con a y b enteros.
   f) Todos los enteros positivos.
   g) Todos los números racionales enteros cuyo denominador sea 1 o una potencia de 2.
5. Hallar dos ejemplos de dominios de integridad no mencionados en el Ejercicio 4.

*6. Probar que la ley conmutativa de la adición es una consecuencia del caso particular $a+(-a)=(-a)+a$, y de las restantes condiciones para un dominio de integridad. (*Sugerencia:* Desarrollar $(a+b)\cdot(1+1)$ de dos modos diferentes, y sumar elementos negativos convenientes.)

3. Propiedades de orden

Otro aspecto de la teoría de los números enteros es la posibilidad de alinearlos en el orden usual:

$$\ldots -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots$$

Este orden es habitualmente expresado con la *relación* $a<b$, entendiendo que $a<b$ (se lee «a es menor que b») equivale a expresar que a está a la izquierda de b en la lista anterior. Pero la relación $a<b$ vale si $b-a$ es un entero positivo, y sólo en este caso. Por consiguiente, cualquier propiedad de la relación $a<b$ podrá ser deducida de las propiedades de los enteros positivos. Consideremos, pues, como nuevos postulados las tres siguientes propiedades de los enteros positivos 1, 2, 3, ... :

Adición: La suma de dos enteros positivos es positiva.

Multiplicación: El producto de dos enteros positivos es positivo.

(*) En lo sucesivo, el asterisco señalará, como ahora, los ejercicios de mayor dificultad.

Ley de tricotomía: Para cualquier entero a resulta válida una, y sólo una, de estas tres alternativas: o es a positivo, o es $a=0$ o es $-a$ positivo.

Estas propiedades las satisfacen también, como es sabido, los números racionales positivos, así como los reales positivos; luego todas las consecuencias de estas propiedades serán también verificadas por tales sistemas de números. Resulta, pues, conveniente llamar *ordenado* a todo dominio de integridad que contenga elementos positivos con las tres propiedades enunciadas.

DEFINICIÓN. *Un dominio de integridad D se dice ordenado si existen en él ciertos elementos, llamados positivos, los cuales satisfacen las leyes de tricotomía, adición y multiplicación, enunciadas arriba para los enteros.*

TEOREMA 2. *En todo dominio ordenado, el cuadrado de cualquier elemento no nulo es positivo.*

Demostración. Sea a^2 el cuadrado dado, con $a \neq 0$. Por la ley de tricotomía, o a o $-a$ será positivo. En el primer caso, a^2 es positivo, por la precedente ley de multiplicación. En el segundo caso, $(-a)^2=a^2$, por la Regla 6 de § 2; luego a^2 es positivo siempre, c. q. d.

Corolario. La unidad $1=1^2$ es siempre positiva.

DEFINICIÓN. *En un dominio ordenado, las dos expresiones* $a<b$ *(se lee «*a *es menor que* b*») y* $b>a$ *(«*b *es mayor que* a*») son equivalentes, y ambas indican que* $b-a$ *es positivo. Además,* $a \leqslant b$ *significa que o es* $a<b$ *o es* $a=b$.

De acuerdo con esta definición, los elementos positivos a se pueden definir como los mayores que cero. Los elementos b menores que cero se llamarán *negativos*. De la misma definición pueden deducirse un gran número de leyes, bastante conocidas, de la relación «menor que».

Ley transitiva: Si es $a<b$ y $b<c$, será $a<c$.

Demostración. Por definición, la hipótesis $a<b$ y $b<c$ significa que $b-a$ y $c-b$ son positivos; luego, por el principio de adición, la suma $(b-a)+(c-b)=c-a$ es positiva, lo cual significa que $a<c$.

Los tres postulados fundamentales de los elementos positivos se reflejan en tres propiedades correspondientes de las desigualdades.

Adición a una desigualdad: Si es $a<b$, será también $a+c<b+c$.

Multiplicación por una desigualdad: Si es $a<b$ y $0<c$, también será $ac<bc$.

Ley de tricotomía: Entre dos elementos cualesquiera a y b es válida una, y sólo una, de las relaciones $a<b$, $a=b$ o $a>b$.

Como ejemplo, vamos a demostrar el principio de multiplicación, según el cual los dos miembros de una desigualdad pueden multiplicarse por un mismo número positivo c. Para ello, deberemos probar que $bc-ac=(b-a)c$ es positivo (cfr. Ejerc. 1 de § 2). Pero esto es consecuencia inmediata del postulado de la multiplicación, pues los factores $b-a$ y c son positivos ambos. Por razonamiento análogo, se demuestra que la multiplicación por un elemento negativo invierte el sentido de la desigualdad (ver el próximo Ejerc. 1 *c*).

Definición. *En un dominio ordenado, se llama valor absoluto* $|a|$ *de un número* a *al cero cuando* a *es cero, y en otro caso, al elemento positivo del par* a, $-a$.

Esta definición se reduce a la siguiente alternativa:

(2) $$|a|=+a \text{ si } a\geqslant 0; \qquad |a|=-a \text{ si } a<0$$

Por consideración separada de cada uno de los casos posibles, se demuestran las leyes del valor absoluto de sumas y productos:

(3) $$|a\cdot b|=|a|\cdot|b|; \qquad |a+b|\leqslant|a|+|b|$$

La ley correspondiente a la suma puede también ser obtenida como sigue: por definición es $-|a|\leqslant a\leqslant|a|$ y $-|b|\leqslant b\leqslant|b|$; sumando estas desigualdades,

$$-(|a|+|b|)\leqslant a+b\leqslant|a|+|b|$$

Esto nos indica que, sea $a+b$ positivo o negativo, su valor absoluto no puede exceder a $|a|+|b|$.

EJERCICIOS

1. Deducir de los postulados de un dominio ordenado, las siguientes reglas:
   a) Si $a<b$ es $a+c<b+c$, e inversamente.
   b) $a-x<a-y$ si, y sólo si, $x>y$.
   c) Si $a<0$ es $ax>ay$ si, y sólo si, $x<y$.
   d) $0<c$ y $ac<bc$ implican que $a<b$.
   e) $x+x+x+x=0$ implica que $x=0$.
   f) $a<b$ implica $a^3<b^3$.

2. Demostrar que la ecuación $x^2+1=0$ no puede tener soluciones en un dominio ordenado.
3. Demostrar diversas leyes relativas a la relación $a \leqslant b$.
4. Demostrar que en todo dominio ordenado $||a|-|b|| \leqslant |a-b|$.
• 5. Demostrar que en un dominio ordenado $a^2=b^2$ implica $a=b$.
• 6. Probar que en un dominio ordenado la ley de simplificación para el producto puede ser deducida de las restantes hipótesis. (*Sugerencia:* Aplicar la ley de tricotomía.)
• 7. Sea D un dominio de integridad en el cual se define una relación $a<b$ que obedece a la ley transitiva y a los principios para la adición y la multiplicación de desigualdades, y a la ley de tricotomía. Probar que eligiendo convenientemente un conjunto de elementos «positivos», D es un conjunto ordenado.

4. Principio de buena ordenación

Los enteros poseen otra propiedad importante, no característica algebraicamente y no compartida por otros sistemas de números. Tal es el

Principio de buena ordenación. Cualquier conjunto de enteros positivos que contenga al menos un elemento, contiene un elemento mínimo.

En otras palabras, cualquier selección dada de enteros positivos contiene un entero particular m tal, que cualquiera que sea el entero a en la selección dada, es $m \leqslant a$. Por ejemplo, el más pequeño entero positivo par es 2. Más generalmente, un conjunto de números se llama *bien ordenado* si cualquiera de sus subconjuntos no vacíos contiene un elemento mínimo; así pues, el principio antedicho indica que los enteros positivos están bien ordenados.

Para destacar la fuerza de este principio, demostremos:

TEOREMA 8. *No hay ningún entero entre 0 y 1.*

Esto se ve inmediatamente sin más que echar una ojeada al orden natural de los enteros, pero lo que pretendemos es probarlo utilizando las hipótesis fundamentales (postulados), sin necesidad de acudir a la referida serie de enteros. Daremos una prueba indirecta. Si hay algún entero c tal que $0<c<1$, la clase C de tales enteros no estará vacía. Por el principio expuesto, hay un entero m mínimo en esta clase; será $0<m<1$. Multiplicando esta desigualdad por el número positivo m resultará $0<m^2<m$. Entonces m^2

es otro entero de la clase C, menor que el supuesto elemento mínimo de C. Esta contradicción demuestra el Teorema 3.

Teorema 4. *Un conjunto S de enteros positivos que incluya al 1 y que incluya al* n+1 *siempre que incluya al* n, *incluye también a cualquier entero positivo.*

Demostración. Bastará probar que el conjunto S′ de todos los enteros no contenidos en S es vacío. Supongamos que S′ no sea vacío; contendrá un elemento mínimo m. Pero $m \neq 1$ por hipótesis; luego, por el Teorema 3, $m > 1$, y $m-1$ deberá ser positivo. Como además $m-1 < m$, resulta que, por la definición de m, $m-1$ debe estar en S. Se deduce por la hipótesis que $(m-1)+1 = m$ estará en S. Esta contradicción demuestra el teorema.

EJERCICIOS

1. Demostrar que para cualquier entero a, $a-1$ es el mayor entero menor que a.
2. ¿Cuáles de los siguientes conjuntos están bien ordenados? a) todos los enteros positivos impares; b) todos los negativos pares; c) todos los enteros mayores que -7; d) todos los enteros impares mayores que 249.
3. Probar que todo subconjunto de un conjunto bien ordenado está bien ordenado.
4. Demostrar que el conjunto de enteros que contiene a -1000 y que contiene a $x+1$, si contiene a x, contiene a todos los enteros positivos.
5. a) Un conjunto S de enteros tiene al entero b como «cota inferior» si $b \leqslant x$ para todo x en S; el mismo b puede pertenecer o no pertenecer a S. Demostrar que cualquier S no vacío que tiene una cota inferior, tiene un elemento mínimo.
   b) Demostrar que cualquier conjunto de enteros no vacío, que tiene una «cota superior», contiene un elemento máximo.

5. Inducción completa. Cálculo con exponentes

Acabamos de formular una completa relación de las propiedades básicas de los enteros, dependiendo de la adición, multiplicación y ordenación. Queda, pues, establecido, para lo sucesivo, que los enteros *forman un dominio de integridad ordenado, J, en el cual los elementos positivos están bien ordenados.* Cualquier otra propiedad de los enteros puede ser probada por un proceso estrictamente lógico, a partir de las que acabamos de postular. En particular, podremos deducir el importante

Principio de inducción completa. Asociemos a cada entero n una proposición $P(n)$, la cual puede ser verdadera o falsa. Si, primero, $P(1)$ es verdadera y, segundo, para cualquier k la verdad de $P(k)$ implica la de $P(k+1)$, entonces $P(n)$ será verdadera para todo entero positivo n.

En efecto, el conjunto de los enteros k para los cuales $P(k)$ es cierta, satisface a las hipótesis del Teorema 4, y, por tanto, a su conclusión.

Como ejemplo de las demostraciones por inducción, estableceremos ahora, para todo dominio de integridad, la ley distributiva general, referente a cualquier número n de sumandos:

$$a(b_1+b_2+\ldots+b_n)=ab_1+ab_2+\ldots+ab_n \tag{4}$$

Una demostración por inducción requiere, primero, la demostración para $n=1$, lo cual es inmediato en este caso. En segundo lugar, supondremos la ley (4) válida para $n=K$, e intentaremos demostrarla para $n=K+1$. Por las leyes asociativas y la distributiva simple, es

$$a(b_1+\ldots+b_{k+1})=a[b_1+\ldots+b_k)+b_{k+1}]=a(b_1+\ldots+b_k)+ab_{k+1}$$

El primer término de este resultado puede desarrollarse según la ley que suponemos válida para K sumandos, luego

$$a(b_1+\ldots+b_{k+1})=(ab_1+\ldots+ab_k)+ab_{k+1}$$

Pero esto es precisamente la ley (4) para K+1 sumandos. Así tenemos una prueba completa, por inducción, de la referida ley.

Por aplicación sucesiva de esta ley podemos probar una ley distributiva más general, en la forma siguiente:

$$\begin{aligned}(a_1+\ldots+a_m)(b_1+\ldots+b_n)=&\\ =a_1(b_1+\ldots+b_n)+\ldots+a_m(b_1+\ldots+b_n)=&\\ =a_1b_1+\ldots+a_1b_n+\ldots+a_mb_1+\ldots+a_mb_n&\end{aligned} \tag{5}$$

Su enunciado es: El producto de dos sumas puede efectuarse sumando todos los productos posibles de un término de la primera suma por un término en la segunda.

Puede también aplicarse el método de inducción al estudio de las potencias con exponentes enteros y positivos, en un dominio

de integridad D. Si n es un entero positivo, la potencia a^n representa el producto $a \cdot a \cdot \ldots \cdot a$ de n factores. Esto mismo puede expresarse con la definición recurrente

(6) $$a^1 = a\,; \qquad a^{n+1} = a^n a \qquad \text{(para todo } a \text{ en D)}$$

la cual hace posible calcular cualquier potencia a^n mediante el cálculo previo de las potencias más bajas. De estas definiciones se pueden deducir las conocidas reglas relativas a exponentes, m y n, números naturales cualesquiera; estas reglas son:

(7) $$a^m a^n = a^{m+n}$$

(8) $$(a^m)^n = a^{mn}, \qquad (ab)^m = a^m \cdot b^m$$

La primera ley, por ejemplo, se demuestra inmediatamente por inducción sobre n. Si $n = 1$, la ley es $a^{n-1}a = a^n$, que es precisamente la definición de a^n. Supongamos ahora que la ley (7) es cierta para cualquier m y para un entero positivo $n = k$, y consideremos la expresión análoga $a^m a^{k+1}$. Será

$$a^m a^{k+1} = a^m(a^k a) = (a^m a^k)a = a^{m+k}a = a^{m+k+1}$$

lo cual ha resultado aplicando sucesivamente la definición, la ley asociativa, la hipótesis inductiva y, por último, otra vez la definición. Lo cual nos prueba la ley (7) para el caso $n = k + 1$, y por tanto, es aplicable la inducción completa.

Frecuentemente se utiliza en las demostraciones el siguiente

Segundo principio de inducción completa. Asociamos a cada número natural n una proposición P(n). Si, para cada m, la hipótesis de que P(k) es verdadera para todo $k < m$ implica la conclusión de que también P(m) es cierta, puede afirmarse que P(n) se cumple para todo n.

En efecto: sea S el conjunto de enteros para los que P(n) es falsa. A menos de ser vacío, tendrá un mínimo m. Por la elección de m, P(k) será verdadera para todo $k < m$; de aquí, por la hipótesis, P(m) debe asimismo ser cierta, obteniendo así una contradicción. La única solución es admitir que S es vacío.

Nota: En el caso $m = 1$, el conjunto de todos los $k < 1$ es vacío, así que en lo anterior se supone implícitamente la demostración directa de P(1).

EJERCICIOS

1. Demostrar por inducción que las siguientes leyes para exponentes positivos son válidas en todo dominio de integridad:

   a) $(a^m)^n = a^{mn}$, b) $(ab)^n = a^n b^n$, c) $1^n = 1$.

2. Demostrar por inducción que $1+2+\ldots+n = n(n+1)/2$.

3. Los coeficientes binómicos $\binom{n}{k}$ se definen por la fórmula

$$\binom{n}{k} = n!/k!(n-k)!, \text{ donde } n! = 1\cdot 2\cdot \ldots \cdot (n-1)n.$$

   a) Demostrar que $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$.

   b) Demostrar por inducción la fórmula del binomio

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

4. Demostrar por inducción que $x_1^2+\ldots+x_n^2 > 0$ excepto si $x_1 = \ldots = x_n = 0$.

5. Demostrar por inducción las siguientes fórmulas sumatorias:

   a) $1+4+9+\ldots+n^2 = n(n+1)(2n+1)/6$

   b) $1+8+27+\ldots+n^3 = [n(n+1)/2]^2$

6. En cualquier dominio ordenado, demostrar que una potencia impar de un elemento negativo es negativa.

*7. Utilizando la inducción, pero no el principio de buena ordenación, demostrar el teorema 3. (*Sugerencia:* Con $P(n)$ signifiquemos $n \geq 1$.)

*8. Utilizando el Ejercicio 7, demostrar el principio de buena ordenación partiendo del de inducción finita. (*Sugerencia:* Sea $P(n)$ la proposición de que cualquier conjunto de enteros positivos con un término $\leq n$ tiene un término inferior.

6. Divisibilidad

Una ecuación $ax = b$ con coeficientes enteros, no siempre tiene solución entera. Cuando existe tal solución, se dice que b es divisible por a.

DEFINICIÓN. *Un entero* b *es divisible por un entero* a *cuando hay algún entero* d *tal, que* b = ad. *Entonces escribimos* a | b; *diremos también que* b *es un múltiplo de* a *y que* a *es un factor o divisor de* b.

He aquí, pues, una nueva relación: «$a \mid b$». Propiedades de ella son las leyes reflexiva y transitiva:

(9) $a \mid a$; $a \mid b$ y $b \mid c$ implica $a \mid c$.

La primera ley (9) es trivial, pues $a=a\cdot 1$ significa $a \mid a$. La segunda hipótesis equivale a decir que $b=ad_1$ y $c=bd_2$, siendo d_1 y d_2 dos enteros; de lo cual resulta $c=a(d_1 d_2)$, o $c \mid a$, c. q. d.

TEOREMA 5. *Los únicos divisores enteros de 1 son ± 1.*

El teorema afirma que si dos enteros a y b son tales que $ab=1$, ha de ser $a=\pm 1$ y $b=\pm 1$. En efecto, $ab=1$ da $|ab|=|a|\cdot|b|=1$. Como son $a\neq 0$ y $b\neq 0$, $|a|$ y $|b|$ son enteros positivos. Como no hay enteros positivos entre 0 y 1 (Teorema 3), por la ley de tricotomía, $|a|\geqslant 1$ y $|b|\geqslant 1$. Si los dos signos, o uno tan sólo, son de desigualdad, el producto $|a|\cdot|b|$ no puede ser igual a 1. Entonces $|a|=1$ y $|b|=1$, y por tanto, $a=\pm 1$ y $b=\pm 1$.

Como $a=a\cdot 1=(-a)(-1)$, todo entero a es divisible por a, $-a$, $+1$ y -1. Los números a y $-a$, por dividirse mutuamente, se llaman «asociados».

DEFINICIÓN. *Dos enteros* a *y* b *se llaman asociados si se verifican las relaciones* a | b *y* b | a. *Los asociados de 1 se llaman unidades.*

Esta definición significa que un entero es una unidad si, y sólo si, es un divisor de 1; con esto, el Teorema 5 establece, simplemente, que las únicas unidades son ± 1. Si a y b son asociados, $a=bd_1$ y $b=ad_2$; luego $a=a(d_1 d_2)$, y, por la ley de simplificación, queda $1=d_1 d_2$. O sea, que d_1 es un divisor de 1 y, por tanto, $d_1=\pm 1$. Por lo tanto, es $b=ad_1=\pm a$, así que los únicos asociados de a son $\pm a$. Dos enteros a y b son asociados si, y sólo si, $|a|=|b|$.

DEFINICIÓN. *Un entero* p *es primo si, siendo distinto de 0 y de ± 1, es divisible únicamente por ± 1 y* $\pm$p.

Los primeros números primos son:

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ \ldots$$

Todo número que no es primo puede descomponerse en un producto de factores primos:

$$128=2^7;\qquad 90=9\cdot 10=2\cdot 5\cdot 3^2;$$
$$672=7\cdot 96=7\cdot 12\cdot 8=7\cdot 3\cdot 2^5$$

Se observa por experiencia que obtenemos los mismos factores primos cualquiera que sea el método de descomposición. Esta unicidad la demostraremos al estudiar el m. c. d.

Llamaremos *factorizar* a la operación de descomponer un número en factores primos.

EJERCICIOS

1. Lista de todos los divisores de 12 y de los de 36.
2. Demostrar que si $a \mid b$ y $a \mid c$, entonces $a \mid (b+c)$.
3. Demostrar: Si b es positivo y no primo, tiene un divisor primo positivo $d \leqslant \sqrt{b}$.
4. Presentar la lista de todos los primos positivos menores de 100. (*Sugerencia:* Suprimir los múltiplos de 2, 3, 5, 7 y usar el Ejercicio 3.)
5. Si $a \mid b$, demostrar que $|a| \leqslant |b|$, cuando es $b \neq 0$.

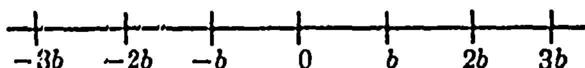
7. El algoritmo de Euclides

El proceso ordinario de dividir un entero a por otro b nos da un cociente q y un resto r. El resultado $a/b = q + r/b$ puede expresarse sin usar explícitamente las fracciones.

ALGORITMO DE LA DIVISIÓN. *Para dos enteros dados* a *y* b, *con* $b > 0$, *existen dos enteros*, q *y* r, *tales que*

$$a = bq + r\,; \qquad 0 \leqslant r < b \tag{10}$$

Imagen geométrica. Si imaginamos los números enteros representados sobre el eje real, los posibles múltiplos bq de b forman un conjunto de puntos equidistantes sobre el eje.



El punto representativo de a debe caer en uno de los intervalos determinados por esos puntos, por ejemplo, en el intervalo bq y $b(q+1)$, excluyendo el punto $b(q+1)$. Esto significa que $a = bq + r$, siendo r menor que la amplitud b del intervalo. Esta imagen sugiere la siguiente demostración, basada sólo en los postulados.

Demostración. Existen ciertamente algunos múltiplos enteros de b que no exceden a a; por ejemplo, como $b > 0$, $b \geqslant 1$ (Teor. 3); así $(-|a|)b \leqslant -|a| \leqslant a$. Por tanto, el conjunto de las diferencias $a - bx$ contiene por lo menos un entero no negativo, a saber, el $a - (-|a|)b$. De aquí, por el postulado de buena ordenación, existe un mínimo no negativo para $a - bx$, al que llamaremos $a - bq = r$. Por construcción, $r \geqslant 0$; mientras que si $r \geqslant b$, entonces $a - b(q+1) =$

$=r-b \geqslant 0$ sería menor que $a-bq$, contra lo afirmado al elegir q. Concluimos, pues, que $0 \leqslant r < b$ y que $a = bq + (a-bq) = bq + r$.

COROLARIO 1. *Dados los dos enteros* a *y* b, *quedan determinados unívocamente el cociente* q *y el resto* r, *que satisfacen a* (10).

Demostración. Supongamos que sea $a = bq + r = bq' + r'$, verificándose $0 \leqslant r < b$ y $0 \leqslant r' < b$. Entonces, $r - r' = b(q' - q)$ es en valor absoluto menor que b, y es múltiplo de b, luego ha de ser 0. De aquí que $r = r'$, $bq = bq'$, $q = q'$.

Frecuentemente deberemos considerar conjuntos de enteros, semejantes al ..., —6, —3, 0, 3, 6, 9, ..., formado por todos los múltiplos de 3. Estos conjuntos tienen la propiedad de que la suma o diferencia de dos cualesquiera de ellos pertenece al conjunto. En general, un conjunto S de números enteros se llama *cerrado* para la adición y la sustracción cuando S contiene la suma $a+b$ y la diferencia $a-b$ de dos enteros cualesquiera, a y b, de S. Todos los enteros pares (positivos, negativos y cero) forman uno de estos conjuntos. Más generalmente, el conjunto de todos los múltiplos xm de un entero fijo m, es cerrado para la adición y sustracción, pues $xm + ym = (x+y)m$ es un múltiplo de m. Ahora vamos a probar que estos conjuntos, constituídos por los múltiplos de un número, son los únicos conjuntos de enteros que tienen dicha propiedad.

TEOREMA 6. *Todo conjunto no vacío de enteros, cerrado para la adición y sustracción, contiene sólo el 0, o contiene un número positivo mínimo, del cual son múltiplos todos los demás.*

Sea S el conjunto, y supongamos que contiene un elemento $a \neq 0$; por definición, S contendrá la diferencia $a - a = 0$, y por tanto, la diferencia $0 - a = -a$; luego S contiene al menos un número positivo, a o $-a$. El principio de buena ordenación nos dice que en S hay un mínimo positivo b.

El conjunto S debe contener todos los múltiplos de b: procediendo por inducción se ve, primero, que contiene $b \cdot 1$ y, segundo, que si está $b \cdot k$ tiene que estar $b \cdot k + b = b(k+1)$. Los múltiplos negativos, tal como $-bn = 0 - bn$ han de estar también, por ser diferencia entre 0 y nb.

Pero S no puede contener enteros no múltiplos de b, pues si hubiera uno a no múltiplo de b, estaría también en S el resto de la división de ambos, $r = a - bq$. Pero r no es negativo y es menor

que b, que es el mínimo entero positivo de S; luego debe ser $r=0$ y $a=bq$.

DEFINICIÓN. *Un entero* d *se llama máximo común divisor (m. c. d.) de dos enteros* a *y* b, *si es simultáneamente divisor de* a *y de* b *y además es múltiplo de cualquier otro divisor común.* En fórmulas, el m. c. d. debe cumplir las tres propiedades siguientes:

$$d\,|\,a\,;\quad d\,|\,b\,;\quad c\,|\,a\ \text{ y }\ c\,|\,b\ \text{ implica }\ c\,|\,d.$$

Por ejemplo, $+3$ y -3 son máximos comunes divisores de 6 y 9. De acuerdo con la definición, si hay varios m. c. d. de dos números, cada uno de ellos debe dividir a los otros; luego serán asociados y diferirán sólo en el signo. Del par $\pm d$ de m. c. divisores de a y b, el número positivo se indicará con el símbolo (a, b).

Nótese que el calificativo «máximo» en la definición de m. c. d. no significa en principio que d tenga mayor magnitud que cualquier otro divisor común c, sino que d es múltiplo de cualquiera de tales c.

TEOREMA 7. *Dos enteros cualesquiera* $a \neq 0$, $b \neq 0$, *tienen un m. c. d. positivo* (a, b). *Este puede expresarse como «combinación lineal» de* a *y* b *con coeficientes enteros* s *y* t, *en la forma*

$$(a, b) = sa + tb \tag{11}$$

Demostración. Consideremos los números de la forma $sa+tb$. Para cada dos

$$(s_1a + t_1b) \pm (s_2a + t_2b) = (s_1 \pm s_2)a + (t_1 \pm t_2)b$$

Por lo tanto, el conjunto S de todos los enteros $sa+tb$ es cerrado para la adición y sustracción y, por el Teorema 6, estará constituído por todos los múltiplos de un número entero positivo $d=sa+tb$. Por esta fórmula, es claro que todo c factor común de a y b ha de serlo de d. Además, los enteros dados, $a=1\cdot a+0\cdot b$, $b=0\cdot a+1\cdot b$, pertenecen ambos a S; luego serán múltiplos del mínimo número d del conjunto. En otras palabras, d es un divisor común al cual dividen todos los demás divisores comunes; luego es $d=(a, b)$, c. q. d.

Análogamente, el conjunto M de los múltiplos *comunes* de a y b es cerrado para la adición y sustracción. Su mínimo elemento

positivo m es un múltiplo común que divide a todos los demás múltiplos comunes. Se le llama mínimo común múltiplo (m. c. m.) de a y b.

TEOREMA 8. *Dos enteros cualesquiera* a *y* b *tienen un mínimo común múltiplo* [a, b], *el cual es divisor de todos los múltiplos comunes, siendo él, a su vez, un múltiplo común.*

Para hallar explícitamente el m. c. d. de dos enteros a y b, se puede utilizar el llamado algoritmo de Euclides. Sean a y b positivos, ya que un entero negativo puede reemplazarse por su asociado positivo sin alterar el m. c. d. [o sea, $(a, b) = (-a, b)$]. El algoritmo de la división da

(12) $$a = bq_1 + r_1, \qquad 0 \leqslant r_1 < b$$

Cualquier entero que divida a los a y b, divide al resto r_1; recíprocamente, todo divisor común de b y r_1 es divisor de a, como resulta por (12). Los divisores comunes del par a, b son, pues, los mismos que los del par b, r_1, así que $(a, b) = (b, r_1)$. Esta reducción puede repetirse con b y r_1, e iterar el proceso

(13)
$$\begin{aligned} b &= r_1 q_2 + r_2, & 0 &< r_2 < r_1; \\ r_1 &= r_2 q_3 + r_3, & 0 &< r_3 < r_2; \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ r_{n-2} &= r_{n-1} q_n + r_n, & 0 &< r_n < r_{n-1}; \\ r_{n-1} &= r_n q_{n+1} \end{aligned}$$

Como el resto disminuye constantemente, habrá finalmente un resto 0, como hemos indicado en la última igualdad. Este razonamiento nos dice que el m. c. d. buscado es

$$(a, b) = (b, r_1) = (r_1, r_2) = \ldots = (r_{n-1}, r_n)$$

Pero la última igualdad de (13) muestra que r_n es divisor de r_{n-1}, así que el m. c. d. de ambos es el propio r_n. El m. c. d. de dos enteros dados a, b, es el último resto distinto de cero que se obtiene aplicándoles el algoritmo de Euclides (12) y (13).

El mismo algoritmo puede utilizarse para representar explícitamente al m. c. d. como combinación lineal $sa + tb$. Esto se consigue expresando los restos sucesivos r_i mediante a y b, en esta forma:

$$\begin{aligned} r_1 &= a - bq_1 = a + (-q_1)b_1, \\ r_2 &= b - q_2 r_1 = (-q_2)a + (1 + q_1 q_2)b \end{aligned}$$

La forma de estas igualdades indica que puede obtenerse r_n como combinación lineal de a y b, con coeficientes enteros, s y t, en cuya expresión intervienen los cocientes q_i.

La forma $(a,b)=sa+tb$ del m.c.d. es de gran utilidad. Una consecuencia importante es que si un número primo divide a un producto de dos factores, debe dividir por lo menos a uno de ellos :

TEOREMA 9. *Si* p *es primo,* p | ab *implica* p | a *o* p | b.

Por definición de número primo, los únicos factores de p son ± 1 y $\pm p$. Si la conclusión $p \mid a$ es falsa, los únicos divisores comunes de p y a son ± 1, así que 1 es un m.c.d. de a y p, y por lo tanto, $1=sa+tp$. Multiplicando por b, resultará :

$$b=sab+tpb$$

Los dos términos de la derecha son divisibles por p, luego b será divisible por p, que es la segunda alternativa del enunciado.

Si $(a,b)=1$, diremos que a y b son primos entre sí. En otras palabras, dos enteros a y b son primos entre sí si no tienen divisores comunes salvo ± 1. La demostración del Teorema 9 prueba también la siguiente generalización :

TEOREMA 10. *Si* (a, c)=*1* *y* c | ab, *debe ser* c | b.

De aquí resulta una consecuencia, relativa a un entero m que sea múltiplo de dos números primos entre sí, a y c. Pues el número m, que es de la forma $m=ad$, es divisible por c, así que, por el teorema, será $c \mid d$, y $m=ad=a(cd')$. Luego el producto ac divide a m. Esto demuestra :

TEOREMA 11. *Supuesto que* (a, c)=*1*, a | m *y* c | m, *se deduce que* ac | m.

EJERCICIOS

1. Mediante el algoritmo de Euclides, calcular el m.c.d. de

a) (14, 35)	b) (11, 15)	c) (180, 252)
d) (2873, 6643)	e) (4148, 7684)	f) (1001, 7655)

2. Escribir (x, y) en la forma $sx+ty$ (s, t enteros), en los tres primeros casos del Ejercicio 1.
3. Demostrar que $(0, a)=|a|$ para cualquier entero a.
4. Si $a>0$, demostrar que $(ab, ac)=a(b, c)$.

5. Demostrar que $b \mid c$ y $|c| < b$, implica $c=0$. (Esto se utiliza en el Corolario 1.)
6. a) Demostrar que tres enteros cualesquiera, a, b, c, tienen un m. c. d. que puede expresarse en la forma $sa+tb+uc$.
   b) Demostrar que $((a, b), c)=(a, (b, c))=((a, c), b)$.
7. Discutir los Ejercicios 3, 5 y 6 b), en el caso del m. c. m.
8. Extender los resultados del ejercicio 6 al m. c. d. de k enteros.
9. Demostrar que el algoritmo de la división vale también para b negativo, si r está limitado por $0 \leqslant r < |b|$.
10. En el algoritmo euclídeo, demostrar por inducción sobre k que cada resto puede expresarse en la forma $r_k=s_k a+t_k b$, siendo s_k y t_k enteros.
11. Dar detalles de la demostración del Teorema 10.
12. ¿Cuáles de los siguientes conjuntos de enteros son cerrados para la adición y sustracción simultáneamente? Cuando se dé este caso, mostrar el mínimo número positivo del conjunto.
    a) Todos los enteros m tales que alguna potencia de m sea divisible por 64.
    b) Todos los m con $(m, 7)=1$.
    c) Todos los m con $m \mid 24$.
    d) Todos los m tales que $6 \mid m$ y $24 \mid m^2$.
    e) Todos los m tales que $21m$ sea divisible por 9.
13. Demostrar que un conjunto de enteros cerrado para la sustracción es forzosamente cerrado para la adición.
14. Demostrar que un conjunto de enteros cerrado para la adición no consiste siempre en los múltiplos de un elemento fijo.
15. Si $1=sa+tb$, para ciertos enteros a, b, s, t, demostrar que a y b son primos entre sí.
16. Si q es un entero tal que, para todos los enteros a y b, $q \mid ab$ implique $q \mid a$ o $q \mid b$, demostrar que q es 0, ± 1 o primo (cfr. Teorema 9).
17. a) Demostrar que si $(a, m)=(b, m)=1$, también $(ab, m)=1$.
    b) Demostrar que si $(a, c)=d$, $a \mid b$ y $c \mid b$, entonces $ac \mid bd$.
    c) Demostrar que $[a, c]=ac/d$.

8. Teorema fundamental de la Aritmética

Ahora resulta fácil demostrar el teorema fundamental de la Aritmética.

TEOREMA 12. *Todo entero distinto de 0 puede expresarse como el producto de* (± 1) *por factores primos positivos. Esta expresión es única, salvo el orden en que los factores se consideren.*

Que todo entero a puede escribirse como un tal producto, puede demostrarse descomponiéndole sucesivamente en factores menores. Este proceso supone el segundo principio de inducción com-

pleta, y puede desarrollarse como sigue (basta considerar enteros positivos):

Sea $P(a)$ la proposición que dice que a puede descomponerse en factores como expresa el enunciado del teorema 12. Si $a=1$ o si a es primo, $P(a)$ es evidentemente cierto. Si a es compuesto, tendrá un divisor positivo b, distinto de 1 y de a, así que $a=bc$, con $b<a$, $c<a$. Pero, de acuerdo con el segundo principio de inducción, podemos suponer que $P(b)$ y $P(c)$ son ciertos, así que b y c pueden expresarse como productos de factores primos:

$$b=p_1p_2\dots p_r, \qquad c=q_1q_2\dots q_s$$

obteniéndose para a la expresión compuesta

$$a=bc=p_1p_2\dots p_rq_1q_2\dots q_s$$

que es de la forma requerida.

Para demostrar la unicidad, consideremos dos posibles descomposiciones en factores primos de un entero a:

$$a=(\pm 1)p_1p_2\dots p_m=(\pm 1)q_1q_2\dots q_n$$

Como todos los números primos p_i y q_j son positivos, las unidades ± 1 de ambas descomposiciones han de ser iguales. El factor p_1 es un divisor de $a=\pm q_1q_2\dots q_n$, así que la aplicación repetida del Teorema 9 asegura que p_1 divide por lo menos a su factor q_j de este producto. Como p_1 divide a q_j y los dos son primos, habrá de ser $p_1=q_j$; ordenado el producto, para que q_j aparezca el primero y simplificando p_1 con q_j queda

$$p_2p_3\dots p_m=q_2'q_3'\dots q_n'$$

donde los acentos indican las q_i en el nuevo orden.

Podemos continuar este proceso hasta que en uno de los dos miembros de la igualdad no quede ningún factor. Tampoco podrán quedar en el otro, así que $m=n$. Hemos, pues, identificado las dos descomposiciones, sin más que reordenar los factores del segundo miembro, como asegurábamos en el teorema de unicidad. En una descomposición puede aparecer un número primo p varias veces. Agrupando los factores iguales, podemos escribir:

(14) $\quad a=\pm p_1^{e_1}p_2^{e_2}\dots p_n^{e_n}$, siendo $(0<p_1<p_2<\dots<p_n)$

El teorema de unicidad demuestra que el exponente e_i correspondiente al factor primo p_i está determinado de modo único para cada entero a.

EJERCICIOS

1. Describir un proceso sistemático para hallar el m. c. d. y el m. c. m. de dos enteros, de los que se conoce la descomposición en factores primos, ilustrándolo con $a=216$, $b=360$, y $a=144$, $b=625$. (*Sugerencia:* Es conveniente usar los exponentes 0 para los factores primos que dividen a uno de los números a o b, pero no al otro.)
2. Si $V_p(a)$ indica el exponente de la más alta potencia del primo p divisor de a, demostrar las fórmulas

1) $$V_p(a+b) \geqslant \text{mín. } \{V_p(a), V_p(b)\};$$
2) $$V_p((a, b)) = \text{mín. } \{V_p(a), V_p(b)\};$$
3) $$V_p(a\ b) = V_p(a) + V_p(b);$$
4) $$V_p([a, b]) = \text{máx. } \{V_p(a), V_p(b)\}.$$

3. Si $\|a\| = 2^{-V_p(a)}$, para V_p como el Ejercicio 2, demostrar que
$$\|ab\| = \|a\| \cdot \|b\| \quad y \quad \|a+b\| \leqslant \text{máx. } (\|a\|, \|b\|)$$
*4. Sea $V(a)$ una función no negativa, con valores enteros, definida para los enteros a que tienen las propiedades 1) y 3) del Ejercicio 2. Demostrar que $V(a)$ es o idénticamente 0 o un múltiplo constante de una de las funciones $V_p(a)$ del Ejercicio 2. (*Sugerencia:* Primero hallar algún p con $V(p) > 0$.)
5. Mediante las fórmulas del Ejercicio 2, demostrar que para enteros positivos a y b, $ab=(a, b)[a, b]$. (Para una segunda demostración, cfr. Ejercicio 17 c), § 7.)
6. Demostrar que el número de primos es infinito (Euclides). (*Sugerencia:* Si $p_1, \ldots, p_n$ son n primos, el producto $p_1 \ldots p_n + 1$ no es divisible por ninguno de estos primos.)
*7. Definir la función $e(n)$ (n entero positivo cualquiera) como el m. c. d. de los exponentes que aparecen en la factorización de n. Demostrar: a) para r y n dados, existe un entero x tal que $x^r = n$ si, y sólo si, $r \mid e(n)$; b) $e(n^r) = r \cdot e(n)$; c) si $e(m) = e(n) = d$, es $d \mid e(m, n)$.
8. Si un producto mn positivo es un cuadrado, y si $(m, n)=1$, demostrar que m y n son ambos cuadrados.
*9. Los posibles triángulos rectángulos con los lados medidos por enteros x, y y z, se pueden hallar como sigue: Supongamos que x, y, z, no tienen factores comunes distintos de ± 1.
   a) Si $x^2+y^2=z^2$, mostrar que x e y no pueden ambos ser impares.
   b) Si y es par, aplicándole el Ejercicio 8, mostrar que $y=2mn$, donde m y n son enteros, $x=m^2-n^2$, $z=m^2+n^2$. (*Sugerencia:* Descomponer z^2-x^2.)

9. Congruencias

Al numerar las horas del día, se acostumbra a contar sólo hasta 12 y volver a empezar. Esta sencilla idea de prescindir de los múltiplos de un número fijo, 12 en este caso, es la base de la noción aritmética de congruencia. Diremos que dos enteros son *congruentes* «módulo 12» si difieren en un entero múltiplo de 12. Por ejemplo, 7 y 19 son congruentes y se escribe $7 \equiv 19$ (mód. 12).

Definición. $a \equiv b$ (*mód.* m) *significa que* $m \mid (a - b)$.

Se puede decir igualmente que $a \equiv b$ (mód. m) cuando la diferencia $a - b$ pertenece al conjunto de los múltiplos de m. Todavía cabe otra definición, basada en que el resto de la división de a por m es único (Corol. 1 de § 7). Podemos, pues, establecer lo que sigue:

Teorema 13. *La condición necesaria y suficiente para que dos enteros* a *y* b *sean congruentes módulo* m, *es que den el mismo resto al dividirlos por* m.

Demostración. Como $a \equiv b$ (mód. m) si, y sólo si, $a \equiv b$ (módulo $-m$), bastará demostrar este teorema en el caso $m > 0$. Supongamos primero que $a \equiv b$ (mód. m). Entonces, $a - b = cm$. Dividiendo b por m, se obtendrá un resto r, $b - mq = r$, $0 \leqslant r < m$. Entonces

$$a = b + cm = (qm + r) + cm = (q + c)m + r$$

Esta ecuación indica que r es el resto de a al dividirlo por m; o sea, que a y b dan el mismo resto.

Recíprocamente, supongamos que el resto es igual y que por ende $a = qm + r$, $b = q'm + r$. En tal caso, $a - b = (q - q')m$ es divisible por m, así que $a \equiv b$ (mód. m).

La relación de congruencia para un módulo fijo m tiene para enteros cualesquiera a, b y c, las siguientes propiedades, que recuerdan propiedades análogas de la igualdad:

Reflexiva:	$a \equiv a$	
Simétrica:	$a \equiv b$ implica $b \equiv a$	para todos: (mód. m)
Transitiva:	$a \equiv b$ y $b \equiv c$ implican $a \equiv c$	

Cada una de estas leyes se demuestra por la definición de congruencia. La ley de simetría así traducida, requiere simplemente que $m \mid (a-b)$ implique $m \mid (b-a)$. La hipótesis es $a-b=dm$, y la conclusión $m \mid (b-a)$, puesto que $b-a=(-d)m$.

La relación de congruencia para un módulo fijo m tiene otra propiedad que también recuerda a las de la igualdad; las sumas y productos de enteros congruentes son también congruentes.

TEOREMA 14. *Si* $a \equiv b$ (*mód.* m), *para todo entero* x *resulta:*

$$a+x \equiv b+x, \quad ax \equiv bx, \quad -a \equiv -b \qquad [\textit{todo}\ (\textit{mód.}\ m)]$$

También aquí la prueba se reduce a recordar la definición. Así, la hipótesis es que $a-b=km$ para algún k; de aquí podemos obtener las conclusiones en la forma

$$m \mid (a+x-b-x), \quad m \mid (ax-bx), \quad m \mid (-a+b).$$

La ley de simplificación, válida en las igualdades, no lo es en las congruencias. Así, $2 \cdot 7 \equiv 2 \cdot 1$ (mód. 12), pero no es $7 \equiv 1$ (módulo 12). Esto sucede por ser 2 divisor del módulo, así que la diferencia $2 \cdot 7 - 2$ será divisible por 12 en tanto se conserve el factor 2. Puede enunciarse la ley de simplificación algo modificada:

TEOREMA 15. *Si* c *es primo con* m,

$$ca \equiv cb\ (\textit{mód.}\ m) \quad \textit{implica} \quad a \equiv b\ (\textit{mód.}\ m)$$

Demostración. De acuerdo con la definición, la hipótesis nos dice que $m \mid (ca-cb)$, o sea, $m \mid c(a-b)$, y por ser m primo con c, por el Teorema 10 resulta $m \mid (a-b)$, esto es, $a \equiv b$ (mód. m), c. q. d.

El estudio de las ecuaciones lineales puede extenderse a las congruencias.

TEOREMA 16. *Si* c *es primo con* m, *la congruencia* $cx \equiv b$ (*módulo* m) *tiene una solución entera* x. *Dos soluciones cualesquiera* x_1 *y* x_2, *son congruentes módulo* m.

Demostración. Por hipótesis, m. c. d. $(c, m)=1$, luego 1 es igual a $sc+tm$ para dos enteros convenientes s y t. Multiplicando por b, $b=bsc+btm$. El último término es múltiplo de m, así que $b \equiv (bs)c$ (mód. m). Esto expresa que $x=bs$ es solución de $b \equiv xc$ (mód. m).

Por otra parte, dos soluciones x_1 y x_2 de esta congruencia, han de dar $cx_1 \equiv cx_2$, por ser la relación de congruencia simétrica y transitiva. Como c es primo con m, se puede simplificar como en el Teorema 15, y resulta $x_1 \equiv x_2$ (mód. m).

Un caso particular importante se presenta cuando el módulo m es primo; entonces, todo entero no divisible por m es primo con él. Esto nos demuestra el siguiente

COROLARIO. *Si* p *es primo y* $c \not\equiv 0$ *(mód.* p*), entonces* $cx \equiv b$ *(mód.* p*) tiene solución única módulo* p.

Consideremos ahora congruencias simultáneas.

TEOREMA 17. *Si los módulos* m_1 *y* m_2 *son primos entre sí, las congruencias*

(15) $$x \equiv b_1 \ (\text{mód. } m_1), \qquad x \equiv b_2 \ (\text{mód. } m_2)$$

tienen una solución común, x. *Dos soluciones cualesquiera son congruentes módulo* m_1m_2.

Demostración. La primera congruencia tiene como solución b_1; la solución más general es $x = b_1 + ym_1$, para cualquier entero y. Esta debe verificar la segunda congruencia $b_1 + ym_1 \equiv b_2$ (mód. m_2) o $ym_1 \equiv b_2 - b_1$ (mód. m_2); como m_1 y m_2 son primos entre sí, podemos resolver esta congruencia por el método del Teorema 16.

Supongamos ahora que x y x' son dos soluciones del sistema (15); será $x' - x \equiv 0$ (mód. m_1) y $x' - x \equiv 0$ (mód. m_2). Como m_1 y m_2 son primos entre sí, la diferencia $x' - x$ es divisible por m_1m_2. Así que $x \equiv x'$ (mód. m_1m_2).

El mismo método de resolución se aplica a dos o más congruencias de la forma $a_ix \equiv b_i$ (mód. m_i), con m. c. d. $(a_i, m_i) = 1$ y con los módulos primos entre sí dos a dos.

EJERCICIOS

1. Resolver las siguientes congruencias:

   a) $3x \equiv 2$ (mód. 5) b) $7x \equiv 4$ (mód. 10)
   c) $243x + 17 \equiv 101$ (mód. 725) d) $4x + 3 \equiv 4$ (mód. 5)
   e) $6x + 3 \equiv 4$ (mód. 10) f) $6x + 3 \equiv 1$ (mód. 10)

2. Demostrar que la relación $a \equiv b$ (mód. m) es réflexiva y transitiva.
3. Demostrar directamente que $a \equiv b$ (mód. m) y $c \equiv d$ (mód. m) implica $a + c \equiv b + d$ (mód. m) y $ac \equiv bd$ (mód. m).

• 4. a) Demostrar que la congruencia $ax \equiv b$ (mód. m) tiene solución si, y sólo si, $(a, m) \mid b$.
b) Demostrar que si $(a, m) \mid b$, la congruencia tiene exactamente (a, m) soluciones incongruentes módulo m. [*Sugerencia:* Dividir a, b y m por (a, m).]
5. Si m es entero, mostrar que $m^2 \equiv 0$ o 1 (mód. 4).
6. Demostrar que $x^2 \equiv 35$ (mód. 100) no tiene solución.
• 7. Demostrar que si $x^2 \equiv n$ (mód. 65) tiene una solución, también tiene solución $x^2 \equiv 65 - n$ (mód. 65). Generalizar este resultado.
8. Si x es un número impar no divisible por 3, demostrar que $x^2 \equiv 1$ (módulo 24).
• 9. a) Mostrar con tablas que todos los números entre 25 y 40 pueden ser expresados como suma de cuatro o menos cuadrados (el resultado es verdadero para todos los números positivos).
b) Demostrar que ningún entero $m \equiv 7$ (mód. 8) puede expresarse como una suma de tres cuadrados. (*Sugerencia:* Generalizar Ejercicio 5.)
10. Resolver las congruencias simultáneas:
a) $x \equiv 2$ (mód. 5) $\quad 3x \equiv 1$ (mód. 8)
b) $3x \equiv 2$ (mód. 5) $\quad 2x \equiv 1$ (mód. 3)
11. En una isla desierta, cinco hombres y un mono recogen cocos durante todo el día, y después se duermen. El primer hombre se despierta y decide tomar su parte. Divide los cocos en cinco grupos iguales, y le sobra un coco, que lo da al mono. Después toma su parte y vuelve a dormirse. Entonces despierta el segundo hombre, y haciendo un montón con los cocos que quedaron, lo divide en cinco partes iguales, y le sobra un coco, que da al mono. Sucesivamente ocurre lo mismo con cada uno de los tres hombres restantes. Encontrar el número mínimo de cocos que formaban el montón original. (*Sugerencia:* Añadir 4 cocos.)
•12. Demostrar por inducción que el Teorema 17 puede generalizarse a n congruencias con módulos primos dos a dos.
•13. Demostrar que si $(m_1, m_2) = (a_1, m_1) = (a_2, m_2) = 1$, entonces las congruencias simultáneas $a_i x \equiv b_i$ (mód. m_i) $(i = 1, 2)$ tienen una solución común, y que dos soluciones cualesquiera son congruentes módulo $m_1 m_2$.
•14. Generalizar el Ejercicio 13 a n congruencias simultáneas.

10. Clases residuales

Desde la más remota antigüedad, el hombre ha distinguido los enteros «pares» 2, 4, 6, ..., de los «impares» 1, 3, 5, ... Las siguientes leyes de cálculo entre pares e impares son también conocidas:

$$\text{(16)} \quad \begin{array}{ll} \text{par} + \text{par} = \text{impar} + \text{impar} = \text{par}, & \text{par} + \text{impar} = \text{impar} \\ \text{par} \cdot \text{par} = \text{par} \cdot \text{impar} = \text{par}, & \text{impar} \cdot \text{impar} = \text{impar} \end{array}$$

Estas igualdades pueden considerarse, no como teoremas relativos a los enteros ordinarios, sino como definición de dos opera-

ciones, «adición» y «multiplicación», en una nueva álgebra de los dos elementos «par» e «impar».

Esta álgebra puede también construirse como un álgebra de restos módulo 2. Los enteros pares son aquellos que divididos por 2 dan resto 0, mientras que los impares dan resto 1. Estos dos restos pueden sumarse y multiplicarse del modo ordinario, cuidando luego de reemplazar el resultado por su resto módulo 2. Esto nos da una tabla :

$$0+0=1+1=0 \qquad 0+1=1$$
$$0\cdot 0=0\cdot 1=0 \qquad 1\cdot 1=1$$

que en esencia es la misma tabla (16) para pares e impares. Inversamente, puede decirse que la igualdad $1+1=0$ es un nuevo modo de escribir la congruencia $1+1\equiv 0$ (mód. 2).

Un álgebra análoga J_n, de n elementos, resultará partiendo de las congruencias módulo n. En la última sección hemos visto que la congruencia tiene las propiedades características de la igualdad ; es reflexiva, simétrica y transitiva, y las congruencias pueden ser multiplicadas y sumadas, como las igualdades. En efecto, el Teorema 14 muestra que si $a\equiv b$ (mód. n) y $c\equiv d$ (mód. n) resulta

$$(17) \qquad a+c\equiv b+d \text{ (mód. } n) \qquad a\cdot c\equiv b\cdot d \text{ (mód. } n)$$

El álgebra J_n de los enteros módulo n se obtiene reemplazando la congruencia módulo n por la igualdad. Según (17), la suma y producto de dos enteros están unívocamente determinados con este nuevo significado de igualdad. Cualquier entero es «igual» a uno de los n restos posibles, 0, 1, 2, ..., $n-1$. Dos de estos restos pueden sumarse (o multiplicarse) en la forma habitual, reduciendo luego el resultado a su resto módulo n, del que viene a ser «igual».

Las tablas para el caso $n=5$ son las siguientes :

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

TEOREMA 18. *En el sistema J_n de los enteros módulo* n, *son válidas para la adición y multiplicación todas las propiedades enumeradas en la definición de un dominio de integridad, excepto la ley de simplificación del producto.*

Demostración. Acabamos de ver que dos elementos cualesquiera de J_n definen unívocamente su suma y su producto. Consideremos ahora la ley distributiva. Como $a(b+c)=ab+ac$ para enteros cualesquiera, se debe tener $a(b+c)\equiv ab+ac$ (mód n), que es la ley distributiva para nuestro nuevo concepto de igualdad en J_n. El mismo tipo de razonamiento se aplica a las otras leyes características de un dominio de integridad, que se expresan mediante identidades entre sumas, productos y elementos negativos. Los primeros miembros de cada identidad son congruentes módulo n con los segundos miembros. Por lo cual las correspondientes expresiones en J_n son iguales.

El único postulado que no se conserva inalterado es la ley de simplificación del producto. Por el Teorema 1, esta ley equivaldría a asegurar la no existencia de divisores de 0 en J_n, así que $ab=0$ debiera implicar o $a=0$ o $b=0$. Pero estas igualdades se traducen en J_n por congruencias entre enteros, de modo que tal ley equivaldría a decir: si $ab\equiv 0$ (mód. n), o es $a\equiv 0$ (mód. n) o es $b\equiv 0$ (mód. n). Esto, a su vez, equivale a decir que $n\,|\,ab$ implica o $n\,|\,a$ o $n\,|\,b$. Pero esta proposición es cierta si n es primo (Teorema 9). Si n no es primo, admite una descomposición $n=ab$ y entonces $n\,|\,ab$ sin que sea $n\,|\,a$ ni $n\,|\,b$. Luego, en este caso J_n no satisfará la ley de simplificación.

TEOREMA 19. *Para que la ley de simplificación para la multiplicación sea válida en J_n, es necesario y suficiente que* n *sea un número primo.*

Hay otros modos más sistemáticos para construir el álgebra de los enteros módulo n. El artificio de reemplazar congruencia por igualdad significa, esencialmente, que todos los enteros que dan el mismo resto en su división por n pueden agruparse, y cada grupo viene a ser un «número» nuevo. Cada uno de tales grupos se llama una «clase residual». Para el módulo 5 hay cinco clases residuales, correspondientes a los posibles restos 0, 1, 2, 3, 4; algunos de estas clases son :

$$1 = \{\ldots, -14, -9, -4, 1, 6, 11, 16, \ldots\}$$
$$2 = \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\}$$
$$3 = \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\}$$

Para cada módulo n, la clase residual r_n determinada por un resto r con $0 \leqslant r < n$, está formada por todos los enteros a que dan el mismo resto r en su división por n. Cada entero pertenece a una, y sólo a una, clase residual, y dos enteros que pertenecen a la misma clase son congruentes mód. n (Teorema 13). Hay n clases residuales módulo n, a saber: $0_n, 1_n, \ldots, (n-1)_n$.

Las operaciones algebraicas en J_n pueden efectuarse directamente sobre estas clases. Supongamos que la suma de dos restos r y s dan en J_n un resto t, o sea $r+s \equiv t$ (mód. n). El mismo resultado obtendríamos si, en vez de tomar los restos r y s, tomásemos otros elementos en las clases correspondientes. Si a está en r_n y b en s_n, entonces $a+b$ está en la clase t_n, que contiene a su suma t, pues $a \equiv r$ y $b \equiv s$ implican $a+b \equiv r+s \equiv t$ (mód. n). En general, el álgebra J_n puede definirse como el álgebra de las clases residuales; para sumar (o multiplicar) dos clases, se eligen dos elementos a y b representativos de estas clases, y se busca la clase residual que contiene a la suma (o al producto) de estos elementos representativos. Si a_n indica la clase residual que contiene a a, esta regla puede formularse así:

$$(a+b)_n = a_n + b_n, \qquad (ab)_n = a_n b_n \tag{18}$$

Por ejemplo, la suma $1_5 + 2_5 = 3_5$, de las clases escritas antes, puede hallarse sumando dos elementos elegidos como representantes de las mismas, $6 + (-13)$ por ejemplo, obteniendo así (-7), que está en la clase 3_5. Otras elecciones, como $-9 + (-3) = -12$, $11 + 7 = 18$, $-14 + 17 = 3$, darán siempre la *misma* suma 3_5.

Las clases residuales que hemos definido mediante los restos, pueden definirse también directamente mediante las congruencias, según el método general que será tratado en el Cap. VI.

EJERCICIOS

1. Construir las tablas de adición y multiplicación para J_5 y J_6.
2. Calcular en J_7: $(3 \cdot 4) \cdot 5$, $3 \cdot (4 \cdot 5)$, $3 \cdot (4+5)$, $3 \cdot 4 + 3 \cdot 5$.
3. Hallar todos los divisores de cero en J_{24} y en J_{21}.

4. Determinar exactamente el conjunto de sumas $x+y$ y productos xy, para x en 4_1, y en 4_1. ¿Cómo están relacionados los conjuntos 4_1+4_1 y $4_1 \cdot 4_1$?
5. Demostrar la ley asociativa para la adición de clases residuales, como en la demostración del Teorema 18.

11. Algunos conceptos básicos de Lógica

Llegados a este punto, conviene nos detengamos a discutir brevemente las nociones fundamentales de igualdad, clase, operación binaria, correspondencia y relación.

La noción de *igualdad* se toma a veces como una de las nociones básicas de Lógica, no sujeta a ulterior definición o análisis matemático. Por el contrario, es a veces conveniente definir una nueva igualdad para elementos de un sistema determinado, tal como hemos hecho en el de los enteros módulo n; debe entonces probarse que la relación de igualdad que se ha definido tiene las propiedades adecuadas. Desde cualquier punto de vista, una noción aceptable de igualdad debe cumplir las tres propiedades siguientes, siendo a, b, c, elementos cualesquiera:

Ley reflexiva:	Para todo a, $a=a$.
Ley de simetría:	Si $a=b$, también es $b=a$.
Ley transitiva:	Si $a=b$ y $b=c$, también es $a=c$.

Las palabras «clase» y «conjunto» se emplean indistintamente para indicar colecciones arbitrarias de objetos; así se hablará de la clase de todos los números impares mayores que 18, o del conjunto de todos los puntos equidistantes de dos puntos fijos. Estos ejemplos ilustran el hecho fundamental de que cualquier *propiedad* determina una *clase*, a saber, la clase de todos los elementos que tienen esa propiedad. Recíprocamente, cada clase C determina una propiedad P_C, pues basta decir que un elemento tiene la propiedad P_C si, y sólo si, pertenece a la clase C. A menudo conviene escribir «$x \in C$» como expresión abreviada de que «x pertenece a la clase C». A veces es útil convenir que el *conjunto vacío* o *nulo*, esto es, sin ningún elemento, sea también un conjunto a considerar.

Dos conjuntos A y B son iguales solamente cuando contienen los mismos elementos, así $A=B$ si cada $x \in A$ es $x \in B$ y recíprocamente. Esta igualdad entre conjuntos tiene las propiedades reflexiva, simétrica y transitiva.

Muchas veces aparecen operaciones sobre *pares* de números, como la adición de dos enteros, la adición de dos clases residuales en J_n, el producto de dos números reales, la sustracción entre dos enteros y otras semejantes. En tales casos hablamos de una «operación binaria». En general, una *operación binaria* «o» sobre un conjunto S de elementos $a, b, c, \ldots$, es una regla que asigna a cada par de elementos a y b de S, un tercer elemento definido unívocamente, $c=a \circ b$, del mismo conjunto S. Asegurar que este resultado, c, está definido unívocamente, implica, en particular, que el elemento $c=a \circ b$ no varía si se reemplazan a o b por elementos iguales de S. En otros términos: $a=a'$ implica:

$$(19) \qquad a \circ b = a' \circ b \quad \text{y} \quad b \circ a = b \circ a'$$

(«ley de sustitución»). Por combinación de las dos partes de (19) se demuestra que

$$(20) \qquad a=a' \text{ y } b=b' \quad \text{implica} \quad a \circ b = a' \circ b'$$

Si la operación $a \circ b$ es de adición, la (20) es simplemente el axioma de Euclides: «iguales sumados con iguales, siguen iguales». Una ley como ésta es, pues, un postulado relativo a la igualdad, que debe ser verificado cuando se introduzca un nuevo tipo de igualdad, como hicimos en el caso del álgebra J_n de los enteros módulo n).

A cada entero a corresponde un valor absoluto $|a|$, que es único, pero cada entero positivo b es el valor absoluto de dos enteros diferentes $\pm b$.

Esto se expresa diciendo que la correspondencia $a \to |a|$ es «pluriunívoca» o, más concretamente, en nuestro ejemplo, que es una correspondencia «dos-uno». Una *correspondencia* $a \to a'$ es cualquier ley que establece para cada elemento de a, de una clase A, un elemento correspondiente a', de otra clase B, se llamará *pluriunívoca* si cada elemento de A tiene un solo elemento correspondiente en B y cada elemento de B corresponde a lo menos un elemento de A. Por ejemplo, he aquí una correspondencia:



entre 25 letras y 9 cifras. Si se prescinde del par $Z \leftrightarrow 0$ se tiene aquí una correspondencia 3 — 1.

Otro ejemplo de correspondencia pluriunívoca lo ofrece la reducción de cada entero a a su clase residual módulo n, esto es : $a \rightarrow a_n$.

Las correspondencias *biunívocas*, llamadas también correspondencias *uno-uno*, son especialmente importantes (*). Una coordinación $a \leftrightarrow a'$ se llama *correspondencia biunívoca* entre un conjunto A y otro conjunto B, si cada elemento a en A tiene en B un correspondiente a', y sólo uno, así como cada elemento b en B es el correspondiente $b = a'$ de un elemento de A, y sólo de uno. Por ejemplo, la correspondencia $x \leftrightarrow x+1$ es una correspondencia biunívoca de enteros con enteros, mientras $x \leftrightarrow 2x$ es una correspondencia biunívoca del conjunto de todos los enteros con el conjunto de los enteros pares. Otro ejemplo familiar se presenta en Geometría analítica : si en un plano se han trazado dos ejes coordenados, x e y, es sabido que los puntos del plano están en correspondencia biunívoca con los pares de números reales (x, y). (**).

Dos enteros cualesquiera pueden estar ligados por «relaciones» muy diversas, como en el caso «$a = b$», «$a < b$», «$a \equiv b$ (mód. 7)» o «$a \mid b$». Cada una de estas frases indica una cierta «relación binaria» entre a y b. Fácilmente pueden mencionarse muchas otras relaciones entre objetos no matemáticos, tal como, entre dos hombres, «es hermano de». Esto es un ejemplo de relaciones existentes fuera de las matemáticas. Para expresar relaciones en general se emplea un símbolo, R, que expresa cualquier relación («R» se sustituye en cada caso por «$=$», o «$<$», o «$|$», etc.). Formalmente, «R» indicará una relación binaria en un conjunto dado S de objetos cuando, dados dos elementos a y b en el conjunto S, o bien a está en la relación R con b (simbólicamente, aRb) o a no está en la relación R con b (en símbolos, $aR'b$).

Son especialmente importantes en un conjunto S las relaciones R que, como sucede con la congruencia y la igualdad, satisfacen a las siguientes leyes :

Reflexiva : aRa para todo a en S.

Simétrica : aRb implica bRa para todos los a, b, en S.

Transitiva : aRb y bRc implica aRc para todos los a, b, c, en S.

(*) La designación *uno-uno* parece preferible; pero adoptamos *biunívoca* por ser ya usual en castellano. (N. del T.)

(**) Véase, además, Cap. XVI, § 2.

Las relaciones que cumplen estas tres propiedades, esto es, las relaciones reflexivas, simétricas y transitivas, se llaman relaciones de *equivalencia*. Por ejemplo, la relación de semejanza (o de congruencia) entre los triángulos de un plano es una relación de equivalencia.

EJERCICIOS

1. ¿Cuáles de las siguientes operaciones binarias $a \circ b$ con enteros a y b son asociativas y cuáles conmutativas? $a-b$, a^2+b^2, $(a+b)/2$, $-a-b$.
2. Determinar cuáles de las propiedades «reflexiva», «simétrica» y «transitiva» son aplicables a cada una de las siguientes relaciones entre los enteros a y b: $a \leqslant b$, $a < b$, $a \mid b$, $a^2+a=b^2+b$, $a < |b|$.
3. Hacer lo mismo con las siguientes relaciones entre humanos: «es padre de», «es hermano de», «es amigo de», «es tío de», «es descendiente de». ¿Debe variarse algo en las respuestas, considerando sólo el conjunto de los varones?

**4. ¿Cómo es la relación «es tío de» respecto a las relaciones «es hermano de» y «es padre de»? ¿Puede establecerse una regla general semejante para deducir una nueva relación de otras dos dadas?

*5. Una relación R se llama «circular» si aRb y bRc implican cRa. Demostrar que una relación es reflexiva y circular si, y sólo si, es reflexiva, simétrica y transitiva.

*6. Mostrar dónde está el error de la siguiente «demostración» de que las leyes simétrica y transitiva para una relación R implican la ley reflexiva: «Por la ley simétrica, aRb implica bRa; por la ley transitiva, aRb y bRa implican aRa».

7. Cada una de las siguientes reglas define una correspondencia entre el conjunto J de los enteros y alguna subclase B de los mismos:

   a) $a \to |a|+1$ b) $a \to a^2$

   c) $a \to 2a+5$ d) $a \to$ m. c. d. $(a, 6)$

   En cada caso, determinar la subclase B y estudiar cuándo la correspondencia es biunívoca entre J y B.
8. Repetir el Ejerc. 7 reemplazando J por la clase J^+ de enteros positivos.
9. ¿Para qué enteros n es la correspondencia $x \to 6x+7$ biunívoca sobre J_n?

12. Sistemas de numeración. Isomorfismo

La notación «decimal» usual (o sistema árabe de numeración) se funda en dividir reiteradamente por 10. Por ejemplo: 327 significa $3(10)^2+2(10)+7$. En el caso general, expresando un número natural k con los símbolos árabes, el último dígito r (última cifra de su expresión decimal) es el resto obtenido cuando k se divide por 10, así que será $k=10q+r$, con $0 \leqslant r \leqslant 9$. Si el cociente q es cero, el símbolo para representar a k consistirá en la sola cifra r;

si $q > 0$, el símbolo para k consistirá en la sucesión de dígitos que representen al entero q, seguidos del nuevo dígito r.

Por otra parte, se puede también demostrar fácilmente, efectuando sucesivas divisiones por 10, que todo entero positivo k puede ser representado *de modo único* en la forma

$$k = (10)^t r_t + (10)^{t-1} r_{t-1} + \ldots + 10^2 r_2 + 10 r_1 + r_0, \qquad (21)$$

en la cual todos los restos $r_0, r_1, \ldots, r_t$, están comprendidos entre 0 y 9 inclusive, y $r \neq 0$. La representación decimal para k es en tal caso la sucesión de cifras $r_t r_{t-1} \ldots r_1 r_0$.

Todas las reglas aprendidas en la enseñanza primaria para sumar y multiplicar enteros positivos (entre ellas, las establecidas para «llevar decenas»), se deducen de las definiciones y las leyes algebraicas de § 2. El cálculo con enteros no positivos no encierra esencialmente nuevas dificultades.

Pero después del estudio de las congruencias, es claro que cualquier entero $n > 1$ podrá proporcionarnos una base de numeración tan lógica como pueda serlo la base 10. Para obtener una notación adecuada, únicamente necesitaríamos adoptar símbolos especiales (cifras) para los enteros 0, 1, 2, ..., $(n-1)$, que pueden aparecer como restos; el método explicado más arriba se puede aplicar en todas sus partes sin más que sustituir el 10 por n (*). En algunos casos, la notación diádica, con base 2, es más útil que el sistema decimal. En esta notación, los primeros enteros son :

Decimal :	1	2	3	4	5	6	7	8	9
Diádica :	1	10	11	100	101	110	111	1000	1001

Se puede establecer y definir *de novo* un sistema de enteros diádicos. Los «enteros diádicos» serían, por definición, el 0 y las sucesiones finitas de cifras 1 y 0 comenzando por 1 y precedidas por un signo + o —. Las sumas y los productos de tales números enteros diádicos se podrían calcular sistemáticamente por las reglas normales ; así, $1001 + 101 = 1110$.

Claramente vemos que los diferentes sistemas de «enteros» que podemos obtener de este modo son equivalentes en un cierto sentido. Para establecer con precisión el significado de esta equiva-

(*) Se ha sugerido a veces la adopción de un sistema duodecimal, con base $n = 12$, empleándose los dígitos α para 10 y β para 11 con el fin práctico de simplificar la expresión de las fracciones 1/3, 1/4, etc. Estos propósitos están en el mismo estado que las propuestas para reformar el calendario o para adoptar un lenguaje universal.

lencia, llegamos a un nuevo y fundamental concepto algebraico: el de *isomorfismo*. El paralelismo observado antes entre la notación decimal y la diádica es una correspondencia biunívoca ante el conjunto $J^{(10)}$ de los enteros decimales (positivos, negativos y cero) y el conjunto $J^{(2)}$ de los enteros diádicos. Además, las reglas para la adición y multiplicación en estos dos sistemas dan resultados correspondientes; la adición de dos enteros decimales y la adición de los dos correspondientes enteros diádicos nos dan resultados que se corresponden en nuestra coordinación. Por ejemplo, $2+5=7 \leftrightarrow 10+101=111$. Por esta razón, la correspondencia entre $J^{(10)}$ y $J^{(2)}$ se dice que es un «isomorfismo».

DEFINICIÓN. *Un isomorfismo entre dos dominios de integridad D y D' es una correspondencia biunívoca* $a \leftrightarrow a'$ *entre los elementos* a *de D y los elementos* a' *de D', que satisface, para cualquier par de elementos* a, b, *a las condiciones*

$$(22) \qquad (a+b)'=a'+b', \qquad (ab)'=a'b'$$

Se dice que los dominios D y D' son isomorfos cuando existe entre ellos tal correspondencia.

De acuerdo con tales leyes, puede decirse que el isomorfismo «conserva sumas y productos». Diciéndolo llanamente, dos dominios de integridad son isomorfos cuando difieren solamente en la notación de sus elementos. Otro ejemplo apropiado es el álgebra de «par» o «impar», comparada con el álgebra de 0 y 1, módulo 2, que fué tratada en el § 10. La correspondencia biunívoca

$$\text{par} \leftrightarrow 0, \qquad \text{impar} \leftrightarrow 1$$

es un isomorfismo entre ambos dominios, ya que los elementos correspondientes se suman y multiplican de acuerdo con las mismas reglas. [Cfr. fórmula (16).]

En cualquier dominio de integridad, la correspondencia idéntica, en la cual cada elemento se corresponde consigo mismo, es (trivialmente) un isomorfismo. Algunos dominios de integridad tienen isomorfismos no triviales consigo mismos. Consideremos, por ejemplo, el dominio $J[\sqrt{3}]$, tratado en § 2, compuesto por el conjunto de números $m+n\sqrt{3}$, con m y n en el dominio J de los enteros; es isomorfo consigo mismo en la correspondencia no trivial

$m+n\sqrt{3} \leftrightarrow m-n\sqrt{3}$. Esta correspondencia es un isomorfismo, ya que para todo $a=m+n\sqrt{3}$ y $b=m_1+n_1\sqrt{3}$ tenemos:

$$\begin{aligned}(ab)' &= [(m+n\sqrt{3})(m_1+n_1\sqrt{3})]' = \\ &= [(mm_1+3nn_1)+(mn_1+m_1n)\sqrt{3}]' = \\ &= (mm_1+3nn_1)-(mn_1-m_1n)\sqrt{3},\end{aligned}$$

$$a'b' = (m-n\sqrt{3})(m_1-n_1\sqrt{3}) = (mm_1+3nn_1)-(mn_1+m_1n)\sqrt{3}$$

y del mismo modo $(a+b)'=a'+b'$.

Un isomorfismo $a \leftrightarrow a'$ no sólo conserva las sumas y productos, sino también las diferencias. Por definición, $a-b$ es la solución de la ecuación $b+x=a$, así que $b+(a-b)=a$; como la correspondencia conserva las sumas, $b'+(a-b)'=a'$; esto asegura que $(a-b)'$ es la única solución de la ecuación $b'+x=a'$, o sea que

$$(a-b)'=a'-b'$$

Otras reglas son:

$$(23) \qquad 0'=0, \qquad 1'=1, \qquad (-a)'=-(a')$$

que se enuncian: el cero y la unidad de D se corresponden respectivamente con el cero y la unidad de D'; y también se corresponden los pares de elementos opuestos.

Más adelante, veremos que el concepto de isomorfismo se aplica a los sistemas algebraicos más generales. Se puede definir el álgebra abstracta como el estudio de las propiedades de los sistemas algebraicos que se conservan en los isomorfismos.

EJERCICIOS

1. Demostrar que cualquier entero k tiene expresión decimal única dada por (21).
2. Efectuar los siguientes cálculos en sistema diádico: $101+1011$, $(111)(101)$, $11(1100+110)$, $(10101)^2+(11)^3$.
3. ¿Cuántas cifras tiene el número 10^3 en el sistema diádico?
4. Establecer la regla general para adicionar dos números en el sistema diádico.
5. ¿Es válida en la base 7 la igualdad $(101)^2=10201$? ¿En qué base lo es? Lo mismo para $(101)\cdot(102)=(10302)$.
6. Un tendero tiene solamente cinco pesas de, respectivamente, 1, 3, 9, 27 y 81 libras. Puede colocar las pesas en ambos platillos de la balanza. Mostrar de qué modo pesará hasta 121 libras.

7. Los cinco primeros múltiplos 18, 27, 36, .. de 9 son tales que la suma de sus cifras es divisible por 9. ¿Es esto cierto en general? ¿Cómo se demuestra?

* 8. Generalizar el resultado del Ejercicio 7 a cualquier base, además de la decimal.

* 9. Dar una regla para extraer la raíz cuadrada en cualquier sistema de numeración, ilustrarla calculando la parte entera de $\sqrt{20\,000}$ en base 7. (*Es decir,* $\sqrt{4\,802}$ *en base decimal.*)

Ejercicios sobre isomorfismos

10. Demostrar que la propiedad (23) es válida para cualquier isomorfismo.

11. Sea $J[\sqrt{2}]$ el dominio de todos los números $m+n\sqrt{2}$, con m, n en J. Mostrar un isomorfismo no trivial de $J[\sqrt{2}]$ consigo mismo.

12. Demostrar que la correspondencia $m+n\sqrt{2} \leftrightarrow m+n\sqrt{3}$ no es un isomorfismo entre los dominios $J[\sqrt{2}]$ y $J[\sqrt{3}]$.

13. a) Demostrar que, bajo cualquier isomorfismo, un elemento x que satisface a la ecuación $x^2=1+1$ debe corresponder a un elemento $y=x'$ satisfaciendo a la ecuación $y^2=1'+1'$.
    b) *Mediante a), demostrar que ningún isomorfismo es posible entre* $J[\sqrt{2}]$ y $J[\sqrt{3}]$.

14. Demostrar que el dominio J de los enteros carece de isomorfismos no triviales consigo mismo.

* 13. Perfección de la axiomática de los enteros

Al *describir el sistema J* de los enteros como un dominio ordenado, en el que cualquier conjunto de enteros positivos tiene un elemento mínimo, señalábamos que estos postulados definían a los enteros para todos los efectos matemáticos. Pero una tal descripción lleva inherente una limitación esencial. Consideremos, en general, dos sistemas S y S', en los que la adición está definida, e imaginemos S isomorfo con S'. Si en S se satisface la ley conmutativa, entonces $a+b=b+a$, para cualesquiera a y b en S. Los elementos correspondientes en el isomorfismo deben ser iguales, y así $(a+b)'=(b+a)'$. Como el isomorfismo conserva las sumas, $a'+b'=b'+a'$. Por lo tanto, la ley conmutativa es válida en S'. La argumentación tiene carácter general y puede aplicarse a todos los postulados. Por lo tanto, ningún sistema de postulados podrá establecer distinción entre dos sistemas isomorfos. Lo más que puede esperarse es demostrar que con determinados postulados se caracteriza a los números enteros «salvo isomorfismos». Por ejemplo, se cumplirán para los enteros expresados en la numeración

decimal normal; pero serán igualmente ciertos para los enteros expresados en la base dos, tres o cualquier otra.

TEOREMA 20. *Cualquier dominio ordenado J' contiene un subdominio isomorfo con el dominio J de los enteros.*

Demostración. El dominio J' tendrá un solo elemento idéntico para la multiplicación, esto es, la unidad 1'. En J' designemos con m' el elemento $1'+\ldots+1'$ (m sumandos), donde m es un número natural. Por la ley general asociativa del § 2, tendremos:

$$\overbrace{(1'+\ldots+1')}^{m \text{ sumandos}}+\overbrace{(1'+\ldots+1')}^{n \text{ sumandos}}=\overbrace{1'+1'+\ldots+1'+1'}^{m+n \text{ sumandos}}$$

y análogamente, por la ley general distributiva (5) de § 5,

$$\overbrace{(1'+\ldots+1')}^{m \text{ sumandos}}\ \overbrace{(1'+\ldots+1')}^{n \text{ sumandos}}=\overbrace{1'\cdot 1'+\ldots+1'\cdot 1'}^{m\cdot n \text{ sumandos}}$$

Como $1'\cdot 1'=1'$, en conclusión resulta, con nuestra notación,

(24) $$m'+n'=(m+n)' \quad \text{y} \quad m'\cdot n'=(mn)'$$

para todos los enteros positivos m y n.

Un razonamiento análogo se aplica al 0 y a los enteros negativos. Designemos por $0'$ el (único) elemento aditivamente idéntico en J'. Por las reglas 1 y 5 de § 2, se obtiene inmediatamente:

(25) $$0'+x=x \quad \text{y} \quad 0'\cdot x=0' \qquad \text{para todo } x \text{ en } J'$$

Denotemos por $(-1)'$ el único x que satisface a $1'+x=0'$ en J'; y sea $(-m)'=(-1)'+\ldots+(-1)'$ (m sumandos). Por la regla 6 del § 2, y repetición del razonamiento que dió (24):

(26) $$(-m)'+(-n)'=[-(m+n)]' \quad \text{y} \quad (-m)'\cdot(-n)'=(mn)'$$

Análogamente, como $1'\cdot(-1)'=(-1)'$, resulta:

(27) $$m'\cdot(-n)'=(-mn)'$$

Finalmente, la expresión de una suma del tipo

$$[1'+\ldots+1'] + [(-1)'+\ldots+(-1)']$$

puede ser simplificada por supresión de los elementos opuestos, suprimiendo igual número de ellos en cada corchete; esto da la regla para sumar positivos con negativos:

(28) $$m' + (-n)' = (m - n)'$$

para $$m > n, \quad m < n \quad \text{o} \quad m = n$$

Como J' es un dominio ordenado, sabemos por § 3 que $1' > 0$, y

(29) $$\ldots < (-2)' < (-1) < 0' < 1' < 2' < \ldots \qquad \text{en } J'$$

y así, en particular, por la ley alternativa,

(30) $$\text{Si } m \neq n, \quad \text{es } m' \neq n' \qquad \text{en } J'$$

Esto significa que la correspondencia $a \rightarrow a'$ es biunívoca entre los elementos $a = \pm m$ de J y al menos un *subconjunto* S de J'. Por (24)-(28) esta correspondencia conserva sumas y productos: es, pues, un isomorfismo. Incidentalmente, vemos por (30) que conserva también el orden (es un «isomorfismo ordenado»). Esto demuestra el Teorema 20.

Probaremos ahora que si J' cumple el principio de buena ordenación, no contiene más elementos que los que acabamos de considerar. Esto implicará el siguiente resultado final:

TEOREMA 21. *Todo dominio ordenado J' en el que los elementos positivos estén bien ordenados, es isomorfo con el dominio J de los enteros.*

Demostración. Por el razonamiento precedente sabemos que J' contiene un subdominio S isomorfo con J en la correspondencia $a' \leftrightarrow a$. Falta sólo probar que cualquier elemento de J' es uno de los elementos a' de S. Supongamos por un momento que no fuese así. Como el opuesto $-m' = (-m)$ de cualquier elemento de S pertenece también a S, podemos suponer que J' contiene algún elemento *positivo* b que no está en S. Como los elementos positivos de J' están bien ordenados, debe existir algún elementos positivo mínimo c que está en J' pero no en S. Aplicando a J' el Teorema 3, se ve que no es posible que $0 < c < 1'$, y $c = 1'$ es imposible porque $1'$ está en S. Por lo tanto, $c > 1'$, así que $c - 1'$ es también un elemento positivo de J. Como éste es menor que c, deberá ser un elemento de S (por la definición de c), lo cual es decir que $c - 1' = m'$, siendo m un elemento de J. Pero entonces $c = m' + 1' = (m + 1)'$,

de modo que c es un elemento de S, como correspondiente del $m+1$ de J. Pero esto es contradictorio con la definición de c; resulta así demostrado que es absurdo suponer la existencia en J′ de elementos no pertenecientes a S.

EJERCICIOS

1. En la demostración del Teorema 20, mostrar que la ley $(m+1)=m'+1'$ vale para cualquier entero m.
2. Demostrar la primera parte del Teorema 20, como sigue: definir 1′ como elemento idéntico de J′, y definir $(m+1)'$, por inducción, como $m'+1$.
   a) Para m dado, indiquemos con P(n) la función $(m+n)'=m'+n'$. Demostrar P(n) para todo n, por inducción.
   b) Análogamente, poniendo Q(n) para $(mn)'=m'n'$. Demostrar Q(n) por inducción.
   c) Demostrar análogamente las leyes (26)-(28).
3. Demostrar que los enteros positivos no están bien ordenados en el dominio de todos los números racionales (fracciones).
4. En el Teorema 20, demostrar que si J′ contiene un elemento que no esté en S, J′ contendrá infinitos elementos positivos que no estén en S.

CAPITULO II

Números racionales y campos

1. Definición de campo

Conocemos ya algunos dominios de integridad en los que es posible la división: esto es, en los que la ecuación $ax=b$ $(a\neq 0)$ tiene siempre solución. Por ejemplo, dicha ecuación es siempre resoluble en el dominio de los números racionales (enteros y fracciones). Ya veremos que es necesario construir fracciones precisamente iguales a las que conocemos por aritmética, para que el dominio J de los enteros pueda sumergirse en otro dominio más amplio, en el que sea posible la división.

Otros dos importantes dominios en los que es posible siempre la división son: el conjunto de todos los números reales y el de todos los números complejos $a+bi$ (en los que a y b son reales e $i^2=-1$). Tales sistemas se llaman campos (*).

DEFINICIÓN. *Un campo* F *es un dominio de integridad que contiene para cada elemento* $\mathrm{a}\neq 0$ *un «inverso»* a^{-1} *que satisface la ecuación* $\mathrm{a}^{-1}\mathrm{a}=1$.

TEOREMA 1. *La división (excepto por cero) es posible y uniforme en todo campo.*

Vamos a demostrar que para todo $a\neq 0$ y b en un campo F la ecuación $ax=b$ tiene solución única, x, en F. Si $a\neq 0$, su in-

(*) Muchos autores llaman *cuerpos* a los sistemas que ahora van a definirse. Los ingleses, en cambio, los denominan siempre campos (fields). Hemos decidido respetar este nombre en nuestra traducción. — N. DEL T.

verso a^{-1} permite construir un elemento $x=a^{-1}b$, el cual, por sustitución directa, se prueba que es solución de $ax=b$. Es única, pues $ax=b$, $ay=b$, implican $ax=ay$ y de aquí $(a^{-1}a)x=(a^{-1}a)y$, de donde $x=y$. La solución de $ax=b$ se denota por b/a (el *cociente* de b por a). En particular, $1/a=a^{-1}$.

Todas las reglas para el cálculo algebraico que han sido expuestas en el Capítulo I, § 2, son satisfechas en los campos, considerados como dominio de integridad. Ahora vamos a establecer que las reglas usuales para el cálculo con cocientes pueden también demostrarse en un campo a partir de los postulados.

TEOREMA 2. *En todo campo, los cocientes obedecen a las siguientes leyes (en las que* $b \neq 0$ *y* $d \neq 0$*):*

1) $(a/b)=(c/d)$ *si y sólo si* $ad=bc$;

2) $(a/b)\pm(c/d)=(ad\pm bc)/(bd)$;

3) $(a/b)(c/d)=(ac/bd)$;

4) $(a/b)+(-a/b)=0$;

5) *si es* $(a/b)\neq 0$, *será* $(a/b)(b/a)=1$.

Demostración de 1). La hipótesis $(a/b)=(c/d)$ significa $ab^{-1}=cd^{-1}$. Esto da:

$$ad=a(b^{-1}b)d=cd^{-1}(bd)=cd^{-1}db=bc.$$

Recíprocamente, si $ad=bc$, entonces

$$a/b=b^{-1}a=b^{-1}add^{-1}=b^{-1}bcd^{-1}=cd^{-1}=c/d.$$

Demostración de 2) Observemos que $x=a/b$ o $y=c/d$ indican las soluciones de $bx=a$ y de $dy=c$. Estas ecuaciones pueden combinarse para dar:

$$dbx=da, \qquad bdy=bc, \qquad bd(x\pm y)=ad\pm bc.$$

Así pues, $x\pm y$ es la única solución $z=(ad\pm bc)/bd$ de la ecuación $bdz=ad\pm bc$.

Demostración de 3). Como antes, las ecuaciones $bx=a$ y $dy=c$ pueden combinarse para dar

$$(bd)(xy)=(bx)(dy)=ac,$$

de la cual sale $xy=(ac)/(bd)$.

Demostración de 4). Sustituyendo en 2) tenemos

$$(a/b)+(-a/b)=(ab-ba)/b^2=0/b^2=0\cdot(b^2)^{-1}=0.$$

Demostración de 5). Sustituyendo en 3) tenemos $(a/b)(b/a)=ab/ba$. Pero ab/ba es la única solución de la ecuación $bax=ab$. Como $x=1$ satisface a esta ecuación, será $ab/ba=1$.

Razonamientos parecidos a los precedentes pueden utilizarse para demostrar otras conocidas leyes, como las siguientes:

(1) $(bd)^{-1}=d^{-1}b^{-1}$, $\quad (-b)^{-1}=-(b^{-1})$,

(2) $a\pm(b/c)=(ac\pm b)/c$, $\quad a(b/c)=ab/c$,

(3) $(a/b)/(c/d)=ad/bc$, $\quad (a/b)/c=a/bc$, $\quad a/1=a$,

(4) $-(a/b)=(-a)/b=a/(-b)$, $\quad (-a)/(-b)=a/b$.

Las demostraciones se dejan al lector como ejercicio.

Si admitimos que los números reales forman un campo, se pueden construir con facilidad otros campos utilizando la noción de subcampo.

DEFINICIÓN. *Un subcampo de un campo dado* F *es un subconjunto de* F *que es asimismo un campo respecto a las operaciones de adición y multiplicación en* F.

Las identidades que se verifican en F (por ejemplo, las leyes conmutativa, asociativa y distributiva) también son ciertas «a fortiori» en cualquier subconjunto de F con tal que puedan ejecutarse en él las operaciones en cuestión. Para probar que un subconjunto S de F es un subcampo, puede prescindirse de los postulados que son identidades y probar únicamente aquellos que implican alguna afirmación de «existencia», tales como la existencia de un inverso. De este modo, se tiene el siguiente resultado:

TEOREMA 3. *Un subconjunto* S *de un campo* F *será un subcampo si* S *contiene el cero y la unidad de* F, *si es cerrado para la adición y la multiplicación y si cada* a *de* S *tiene su opuesto* $-$a *y (siempre que sea* a $\neq 0$*) su inverso* a^{-1} *en* S.

Este teorema va a aplicarse ahora para probar que el conjunto de todos los números reales de la forma $a+b\sqrt{2}$, con coeficientes racionales a y b, es un subcampo del campo de los números reales. Este subcampo se indica por $R(\sqrt{2})$, donde R designa el cam-

po de los racionales. El teorema 3 es aplicable, pues la suma de dos números de $R(\sqrt{2})$ es otro número de $R(\sqrt{2})$ y semejantemente el producto es

$$(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(bc+ad)\sqrt{2}.$$

Además, $R(\sqrt{2})$ contiene $0=0+0\sqrt{2}$, $1=1+0\sqrt{2}$ y $-(a+b\sqrt{2})=-a-b\sqrt{2}$. Finalmente, el inverso $(a+b\sqrt{2})^{-1}$ de un elemento distinto de cero puede hallarse racionalizado, el denominador

$$\frac{1}{a+b\sqrt{2}}=\frac{1}{a+b\sqrt{2}}\left(\frac{a-b\sqrt{2}}{a-b\sqrt{2}}\right)=\left(\frac{a}{a^2-2b^2}\right)-\left(\frac{b}{a^2-2b^2}\right)\sqrt{2}.$$

El nuevo denominador a^2-2b^2 no es nunca cero (como se probará en el Capítulo III) y el número inverso que resulta, tiene la forma requerida $a'+b'\sqrt{2}$ con coeficientes racionales, $a'=a/(a^2-2b^2)$, $b'=b/(a^2-2b^2)$. Puede verse que este inverso satisface idénticamente a la ecuación $(a'+b'\sqrt{2})(a+b\sqrt{2})=1$.

De modo análogo, el conjunto $R(\sqrt[3]{5})$ de todos los números reales $a+b\sqrt[3]{5}+c\sqrt[3]{25}$ con coeficientes racionales es un campo. La adición, sustracción y multiplicación se efectúan dentro de este conjunto como en $R(\sqrt{2})$, empleando esta vez el hecho de que $(\sqrt[3]{5})^3=5$ es un número racional. Finalmente, $(a+b\sqrt[3]{5}+c\sqrt[3]{25})^{-1}$ puede calcularse observando que la ecuación

$$(a+b\sqrt[3]{5}+c\sqrt[3]{25})(x+y\sqrt[3]{5}+z\sqrt[3]{25})=1+0\sqrt[3]{5}+0\sqrt[3]{25}$$

es equivalente a un sistema de ecuaciones lineales simultáneas. Estas ecuaciones pueden resolverse para x, y, z, a menos que sean $a=b=c=0$. (Cfr. además Capítulo XIV, § 2.)

Podemos construir todavía otros subcampos, partiendo de que existe un campo de números complejos $a+bi$, en el que $i=\sqrt{-1}$, y a y b son reales. La ecuación cuadrática

$$\omega^2+\omega+1=0$$

tendrá una raíz $\omega=(-1+i\sqrt{3})/2=(-1/2)+(\sqrt{3}/2)i$ en tal campo. (Notemos que como $\omega^3-1=(\omega-1)(\omega^2+\omega+1)=0$, ω es una raíz cúbica imaginaria de la unidad). Todo $a+b\omega$ (a, b racionales) forma un subcampo $R(\omega)$ del campo de los números complejos. Pues, en efecto,

$$(a+b\omega)+(c+d\omega)=(a+c)+(b+d)\omega,$$
$$(a+b\omega)(c+d\omega)=ac+(bc+ad)\omega+bd\omega^2=(ac-bd)+(bc+ad-bd)\omega,$$

habiéndose utilizado la igualdad $\omega^2 = -\omega - 1$ para hacer desaparecer la potencia ω^2. Además, cada $a + b\omega \neq 0$ tiene un inverso en el conjunto, ya que

$$(a+b\omega)\left[\frac{-(b-a+b\omega)}{a^2-ab+b^2}\right] = \frac{a^2-ab+b^2}{a^2-ab+b^2} = 1.$$

El denominador $a^2 - ab + b^2$ que aparece en el inverso no es nunca cero, pues $a^2 - ab + b^2 = (a^2+b^2)/2 + (a-b)^2/2$ es siempre positivo, a menos que $a = b = 0$.

Por analogía con la noción de subcampo, un subconjunto de un dominio de integridad D es denominado un *subdominio* de D, si es asimismo un dominio de integridad con la adición y multiplicación de D. Así, los enteros forman un subdominio del campo de los racionales. Para caracterizar los subdominios, existe un teorema análogo al 3 (omitiendo la existencia de inverso a^{-1}).

EJERCICIOS

1. Dar una definición de campo en que no intervenga la expresión «dominio de integridad», partiendo de las nociones previas.
2. Admitiendo que el conjunto de números reales sea un campo, ¿cuáles de los siguientes conjuntos son subcampos: a) todos los enteros positivos; b) todos los números $a+b\sqrt{3}$ con a, b racionales; c) todos los números $a+b\sqrt[3]{5}$ con a y b racionales; d) todos los números racionales no enteros; e) todos los números $a+b\sqrt{5}$ con a y b racionales, y f) todos los números complejos $a+b\sqrt{2}i$ con a, b racionales (supóngase que los números complejos son un campo)?
3. ¿Forma un campo el conjunto de todos los polinomios con coeficientes reales? ¿Y un dominio de integridad?
4. Demostrar que en el Teorema 3 las condiciones $0 \in S$ y $1 \in S$ pueden reemplazarse por las condiciones «S contiene al menos dos elementos». (*Sugerencia:* $aa^{-1} = 1$.)
5. Demostrar las fórmulas (1)-(4) a partir de los postulados de un campo. Especificar qué elementos deben suponerse distintos de 0 en estas fórmulas.
6. Demostrar la ley de simplificación para la multiplicación a partir de los otros postulados del campo. (*Sugerencia:* Ver la demostración de la regla 2 en Cap. I, § 2.)
7. ¿Un dominio de integridad isomorfo a un campo es también un campo? ¿Por qué?
8. Demostrar que el único subcampo del campo R de los números racionales es el mismo R.
9. Establecer y probar un teorema análogo al 3 para los subdominios.

10. Demostrar que un subcampo del $R(\sqrt{2})$ que contiene a todos los números racionales es el mismo R o el campo completo $R(\sqrt{2})$.
11. Si S y S' son dos subcampos de un campo dado F, demostrar que el conjunto de elementos comunes a S y S' es un subcampo.
12. Establecer un teorema general sobre los subdominios posibles de J. Lo mismo para J_n.
*13. Construir las tablas de adición y multiplicación para un campo de cuatro elementos, suponiendo que $1+1=0$ (la adición es módulo 2), y que hay un elemento x tal que $x^2=x+1$.
*14. Hallar todos los subcampos del campo del ejercicio 13.

2. Construcción de los elementos racionales

Los sistemas considerados en el § 1 son campos, según hemos visto. Sin embargo, no lo hemos probado partiendo de los postulados del Capítulo I, sino admitiendo que los números reales (y complejos), tal como los conocemos por la aritmética, forman un campo. Este proceder no está de acuerdo con el programa que nos hemos trazado, de basarnos únicamente en los postulados para los enteros, enunciados en el Capítulo I.

Al objeto de continuar con este programa, vamos a definir los números racionales a partir de los enteros, y probar, tomando como base los referidos postulados para éstos, que tal definición engendra un campo. Más tarde (Capítulo III) definiremos los números reales mediante los racionales y probaremos (al menos en líneas generales) que los números reales forman un campo que contiene $\sqrt{2}$, $\sqrt[3]{5}$, etc.

En el Capítulo V definiremos los números complejos partiendo de los reales, y probaremos (basándonos en nuestros postulados) que constituyen un campo que contiene a $\sqrt{-1}$. Finalmente, en el Capítulo XIV veremos que esta construcción de los números complejos partiendo de los reales tiene como análoga la construcción directa (o sea, sin mencionar para nada los números reales) de los campos $R(\sqrt{2})$ y $R(\sqrt[3]{5})$ del § 1, partiendo del campo de los números racionales.

Los enteros solos no forman un campo; la construcción de los números racionales a partir de los enteros es esencialmente la construcción de un campo que contenga a los enteros. Naturalmente, este campo deberá, además, contener las soluciones de todas las ecuaciones $bx=a$ con coeficientes enteros a y $b\neq 0$. La construcción abstracta de «números racionales» que resuelvan estas ecua-

ciones se consigue, simplemente, introduciendo ciertos símbolos nuevos $r=(a, b)$, a los que llamaremos *pares*, cada uno de los cuales es solución de una ecuación $bx=a$. Debemos ahora hacer ver que estos nuevos entes pueden sumarse, multiplicarse e igualarse exactamente como los cocientes a/b en un campo (Teorema 2, reglas I-III).

DEFINICIÓN. *El conjunto* R *de números racionales está constituido por todos los pares* (a, b) *de enteros* a *y* $b \neq 0$. *La «igualdad» de pares se rige por el convenio de que*

(5) $(a, b) \equiv (a', b')$ *si, y sólo si,* $ab' = a'b$

mientras que las sumas y productos se definen así:

(6) $(a, b) + (a', b') = (ab' + a'b, bb'),$

(7) $(a, b) \cdot (a', b') = (aa', bb').$

Los resultados son siempre pares con un segundo elemento $bb' \neq 0$.

Pretendemos ahora establecer como igualdad la relación «$\equiv$» de «congruencia» entre pares. Como esta relación no es una identidad formal [(a, b) idéntico a (a', b') significaría $a=a'$, $b=b'$], deberemos probar que la relación tiene las propiedades de la igualdad expuestas en el § 11 del Capítulo I (para la identidad formal, estas propiedades serían triviales). Primeramente podemos comprobar de modo directo que la relación «$\equiv$» es reflexiva, simétrica y transitiva. Después comprobaremos que la suma y el producto están determinados unívocamente según dicha relación. Por ejemplo, $(a, b) \equiv (a', b')$ implica que $(a, b) + (a'', b'') \equiv (a', b') + (a'', b'')$. Pues en efecto, cada suma de la conclusión está definida por una fórmula como la (6), y los dos resultados serán congruentes en el sentido de (5) solamente si se verifica que $(ab'' + a''b)b'b'' = (a'b'' + a''b')bb''$. Pero esta igualdad se deduce de la hipótesis $(a, b) \equiv (a', b')$ (es decir, $ab' = a'b$). Una unicidad análoga vale para la multiplicación. De todo ello podemos concluir que la igualdad definida por (5) tiene las propiedades requeridas.

Pueden probarse ahora varias leyes algebraicas para los números racionales que hemos definido. Así, en la ley distributiva, se pueden reducir sistemáticamente ambos miembros de la igualdad

de acuerdo con las definiciones (6) y (7) del siguiente modo (suponemos que r, r' y r'' son tres pares):

$r(r'+r'')$	$rr'+rr''$
$(a, b)[(a', b')+(a'', b'')]$	$(a, b)(a', b')+(a, b)(a'', b'')$
$(a, b)(a'b''+a''b', b'b'')$	$(aa', bb')+(aa'', bb'')$
$(aa'b''+aa''b', bb'b'')$	$(aa'bb''+aa''bb', bb'bb'')$

Estos dos resultados dan pares iguales según (5), ya que el segundo resultado difiere del primero sólo en la presencia de un factor b en todos los términos. Pero un factor extra en un par da siempre otro par igual, $(bx, by)\equiv(x, y)$, pues por (5) esta igualdad se refiere simplemente a la identidad $bxy=byx$.

Esta demostración explícita de la ley distributiva para números racionales (o pares) es sólo un ejemplo del método. Por el mismo empleo directo de las definiciones y de las leyes en los enteros, se prueban la conmutativa y la asociativa. Un elemento idéntico para la adición (un cero) es el par (0, 1), ya que

$$(0, 1)+(a, b)=(0\cdot b+1\cdot a, 1\cdot b)=(a, b).$$

La ley de simplificación se conserva y el par (1, 1) es elemento idéntico para la multiplicación. El opuesto de (a, b) es $-(a, b)=(-a, b)$. Se cumplen, pues, todos los postulados que definen un dominio de integridad, enunciados en el § 2 del Capítulo I.

TEOREMA 4. *El conjunto* R *de números racionales* r, *construído por pares de enteros, es un dominio de integridad en el que cada ecuación* $rx=1$ *con* $r\neq 0$ *tiene una solución* x *en* R.

Solamente nos queda por demostrar la última afirmación, o sea la existencia de un inverso para cada r. Podemos demostrar, más generalmente, la posibilidad de la división, pues la ecuación

(8) $$(a, b)(x, y)\equiv(c, d) \quad \text{con} \quad (a, b)\not\equiv(0, 1)$$

tiene la solución sugerida por (3),

(9) $$(x, y)=(bc, ad).$$

En efecto: por sustitución directa resulta, $(a, b)(bc, ad)=(abc, bad)$, pero $(abc, bad)\equiv(c, d)$ por ser $abcd=badc$. Por hipótesis $(a, b)\not\equiv(0, 1)$, luego es $a\neq 0$; por tanto, (x, y) tiene un segundo tér-

mino ad no nulo, como se exige en nuestra definición de número racional.

Este nuevo dominio de números racionales así construído a partir del de los enteros, será la ampliación deseada de este último, si se logra que los números enteros aparezcan entre los racionales. Pero esto no es estrictamente posible, porque un par de enteros no puede ser la misma cosa que un entero. Sin embargo, cada entero determina un par $(a, 1)$ que se comporta ante la adición y la multiplicación, exactamente igual que el entero a, como se muestra así:

$$(a, 1)+(b, 1)=(a\cdot 1+b\cdot 1, 1\cdot 1)=(a+b, 1)$$
$$(a, 1)(b, 1)=(ab, 1\cdot 1)=(ab, 1).$$

Puede, pues, afirmarse que la correspondencia biunívoca $a \leftrightarrow (a, 1)$ es un isomorfismo entre el primitivo dominio J de los enteros y un subconjunto del dominio de los racionales. Pero ya sabemos que los postulados definen a los enteros no unívocamente, sino a menos de un isomorfismo (Teorema 21 en el Capítulo I), así que ninguna de sus propiedades algebraicas se pierde si cada entero a se *identifica* con el correspondiente par $(a, 1)$. Bajo este convenio, las ecuaciones (8) y (9) indican que cada par $r=(a, b)$ es la solución de una ecuación $(b, 1)r=(a, 1)$ o $br=a$; de aquí que $r=(a, b)$ es el cociente a/b. El resultado puede enunciarse así:

TEOREMA 5. *El dominio* J *puede sumergirse en un campo* R, *siendo cada elemento de* R *un cociente de enteros de* J.

Los números racionales de R están caracterizados precisamente por el modo como hacen posible la división, en la forma siguiente:

TEOREMA 6. *Si el dominio* J *de los enteros está sumergido en un campo* F, *el conjunto* R′ *de todos aquellos elementos de* F *que son cocientes de enteros, es un campo isomorfo al campo* R *de los números racionales, bajo una correspondencia (de* R *a* R′*) en la que cada entero (en* R*) se corresponde consigo mismo (en* R′*).*

Nota. Un isomorfismo entre dos campos F y F' significa, simplemente, un isomorfismo entre F y F' considerados como dominios de integridad. Concretamente, es una correspondencia biunívoca entre F y F' tal, que si $x \leftrightarrow x'$ e $y \leftrightarrow y'$, debe ser

$$(x+y) \leftrightarrow (x'+y') \quad \text{y} \quad (xy) \leftrightarrow (x'y').$$

Demostración. El campo F contiene los cocientes a/b, que son solución de las respectivas ecuaciones $bx=a$ con coeficientes enteros a y $b\neq 0$. El conjunto R' de estos cocientes contiene a todos los enteros, $a/1=a$; según las reglas del Teorema 2, R' será cerrado para la adición, sustracción, multiplicación y división, así que R' puede ser considerado como el cierre de J para estas operaciones en F. En todo caso, R' es un campo (Teorema 3).

La manera como estos cocientes a/b se suman, multiplican e igualan es la establecida en las reglas I-III del Teorema 2. Exactamente las mismas reglas se utilizan para los pares (a, b) (números racionales). Así pues, la correspondencia $a/b \leftrightarrow (a, b)$ es un isomorfismo entre el cierre R' de J y los racionales R. Obsérvese, en particular, que en esta correspondencia, la imagen de cada entero a de J es $a/1 \leftrightarrow (a, 1)=a$.

La construcción de los números racionales se ha hecho a partir del conjunto J de enteros utilizando solamente las propiedades que hacen de J un dominio de integridad. Así, el hecho de que los enteros constituyan un conjunto ordenado, no ha intervenido en las demostraciones. De aquí resulta que podemos obtener una generalización abstracta de los Teoremas 5 y 6.

TEOREMA 7. *Para un dominio de integridad dado, existe y es única la ampliación a un campo mínimo.*

Este enunciado, detallado algo más, nos dice que cualquier dominio de integridad D puede ser sumergido en un campo F constituído por pares de elementos de D, al que llamaremos el *campo de cocientes* de D (cada elemento de F es un cociente a/b de elementos a y b de D). Cualquier otro campo F' que contenga a D, contiene un subcampo F_0' isomorfo con el campo F de cocientes de D. En este isomorfismo, cada elemento de D se corresponde consigo mismo.

EJERCICIOS

1. Demostrar detalladamente las leyes conmutativa y asociativa para la multiplicación de pares.
2. Demostrar que la relación «igualdad» definida por (5) es reflexiva, simétrica y transitiva.
3. Sea $J[i]$ el conjunto de todos los números complejos $a+bi$, con a y b enteros e $i^2=-1$. a) Mostrar explícitamente cómo se suman y multipli-

can dos de tales números. b) Demostrar que forman un dominio de integridad. c) Hallar su campo de cocientes.

4. ¿Puede el sistema J_6 de los enteros módulo 6 ser sumergido en un campo? ¿Por qué?
5. Describir el campo de cocientes del sistema J_5 de los enteros módulo 5.
6. ¿Cuál es el campo de cocientes del campo R? Generalizar.
7. Demostrar que bajo cualquier isomorfismo $F \leftrightarrow F'$ entre dos campos, $a \leftrightarrow a'$, $b \leftrightarrow b'$ y $c \leftrightarrow c'$, implica $c^{-1} \leftrightarrow c'^{-1}$ y $(a-b)/c \leftrightarrow (a'-b')/c'$, supuesto $c \neq 0$. (Cfr. Ejercicio 10 del Capítulo I, § 12.)
8. Demostrar que la correspondencia $a+b\sqrt{7} \leftrightarrow a+b\sqrt{11}$ (a, b racionales) no es un isomorfismo.
*9. Demostrar que no hay ningún isomorfismo entre el campo $R(\sqrt{7})$ de los números de la forma $a+b\sqrt{7}$ y los de la forma $a+b\sqrt{11}$ (a, b racionales). (*Sugerencia:* Mostrar que no hay correspondiente a $\sqrt{7}$.)
10. ¿Qué puede decirse sobre los campos de cocientes a/b y a'/b' de dos dominios de integridad isomorfos D y D'? Demostrar la afirmación.
*11. Demostrar que cualquier número racional distinto de 0 o ± 1 puede expresarse unívocamente en la forma $(\pm 1)p_1^{e_1} \ldots p_r^{e_r}$, donde las p_i son positivos primos distintos y los exponentes e_i son enteros positivos o negativos.
*12. Demostrar que cualquier número racional, $r/s \neq 0$, puede expresarse unívocamente en la forma $r/s = b_1 + b_2/2! + b_3/3! + \ldots + b_n/n!$, donde n es un entero conveniente, y cada b_k es entero, con $0 \leqslant b_k < k$ si $k > 1$ y $b_n \neq 0$.
13. Para un primo p dado, mostrar que el conjunto $J_{(p)}$ de todos los racionales m/n, con n primo con p, es un dominio de integridad. Identificar su campo de cocientes.
14. Hallar el menor dominio de integridad contenido en R y conteniendo a los números racionales 1/6 y 1/5.
*15. Describir todos los dominios de integridad posibles que sean subdominios de R.

3. Congruencias simultáneas de varias variables

Un campo no consta necesariamente de una infinidad de números; por ejemplo, si p es un número primo, los enteros módulo p forman un campo que sólo contiene un número finito de elementos distintos (esto es, incongruentes). El hecho de que el dominio J_p constituya un campo, es corolario del siguiente Teorema:

TEOREMA 8. *Todo dominio de integridad finito es un campo.*

La hipótesis de que D es finito significa que los elementos de D pueden ser completamente enumerados en una sucesión $b_1, \ldots, b_n$, siendo n un entero positivo (sobre los conjuntos finitos en general, ver el Capítulo XII). Para probar que D es un campo, bastará

probar que cada elemento $a \neq 0$ en D tiene un inverso a^{-1} también en D. Efectuemos todos los productos

(10) $ab_1, \ldots, ab_n$ ($b_1, \ldots, b_n$ son los elementos de D).

Estos productos son elementos de D todos diferentes, pues $ab_i = ab_j$ daría, por la ley de simplificación, $b_i = b_j$, contra la hipótesis de que las b son distintas. Como estos productos (10) han de ser los mismos elementos de D, uno de ellos será la unidad, o sea, $1 = ab_i$. Por lo tanto, el elemento b_i es el inverso de a.

Según esta demostración, para hallar el inverso de a en J_p sería necesario ensayar todos los elementos b_i multiplicándolos por a y viendo si el producto es 1. Pero puede procederse directamente, pues la ecuación $ax = 1$ en J_p con $a \neq 0$ es otra forma de la congruencia $ax \equiv 1$ (mód. p) con $a \neq 0$, y ésta puede ser resuelta calculando el entero x mediante el algoritmo de Euclides, como se vió en el Teorema 16 del Capítulo I. Dicho sea de paso, el corolario de este teorema da una prueba directa del Teorema 8 para el caso del dominio J_p.

De la consideración de una sola ecuación $ax = b$, podemos pasar ahora a los sistemas de ecuaciones simultáneas. Estos sistemas aparecen en el estudio de las redes de corriente eléctrica, al tratar la intersección do planos en el espacio, en la resolución de ecuaciones diferenciales por coeficientes indeterminados y en muchas otras cuestiones. Un conjunto de m ecuaciones lineales simultáneas con n incógnitas $x_1, \ldots, x_n$ puede escribirse en la forma

$$\text{(E)}\qquad \begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}$$

Con notación más breve escribiremos únicamente la i-ésima ecuación indicando su forma por un simple término $a_{ij}x_j$, y la notación Σ que expresa la suma de todos los términos para $j = 1, 2, \ldots, n$:

$$\text{(E)}\qquad \sum_{j=1}^{n} a_{ij}x_j = b_i \qquad (i = 1, 2, \ldots, m)$$

El proceso de eliminación de incógnitas que puede aplicarse a estas ecuaciones, se enseña ordinariamente en el Algebra elemental, donde se supone que los coeficientes son números ordinarios

(racionales o reales). Sin embargo, como tales métodos se basan sólo en las reglas para las operaciones racionales, pueden aplicarse a cualquier campo. En particular, serán aplicables en el campo de los enteros módulo p, cuando el sistema dado, (E), consista en ecuaciones simultáneas de congruencia (mód. p) con coeficientes enteros. Con esta generalización vemos valorizada la definición abstracta de campo.

El proceso de eliminación para un sistema (E), con coeficientes a_{ij} y b_i en cualquier campo F, puede especificarse como sigue: Si todos los coeficientes a_{ij} son cero, las ecuaciones son triviales; no tienen solución a menos que sean todos los $b_i=0$, y en este caso cualquier conjunto de x_j en F será solución. Si algún coeficiente a_{ij} es distinto de cero, podemos ordenar las ecuaciones, y dentro de ellas las incógnitas, de modo que el coeficiente $a_{11}\neq 0$. Dividiremos la primera ecuación por a_{11} para obtener una ecuación equivalente en la cual el coeficiente de x_1 es 1. La variable x_1 puede entonces ser eliminada de las restantes ecuaciones por el proceso de restar a_{k1} veces la primera ecuación de la k-ésima, para $k=2$, 3, ..., n. Se obtiene así un nuevo sistema de ecuaciones de la forma

$$
(E') \qquad
\begin{array}{r}
x_1+c_{12}x_2+\ldots+c_{1n}x_n=b_1' \\
c_{22}x_2+\ldots+c_{2n}x_n=b_2' \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
c_{m2}x_2+\ldots+c_{mn}x_n=b_m'
\end{array}
$$

Cada ecuación de (E') se ha obtenido de dos ecuaciones de (E). Recíprocamente, la k-ésima ecuación del sistema original puede deducirse de nuevo de (E') sumando a_{k1} veces la primera ecuación de (E') a su ecuación k-ésima. El sistema (E') es, pues, equivalente al (E), en el sentido de que cada conjunto de elementos x_1, ..., x_n de F que satisfaga a (E') deberá satisfacer a (E), y viceversa.

En las $m-1$ últimas ecuaciones de (E') intervienen solamente las variables x_2, ..., x_n y pueden reducirse nuevamente mediante el mismo proceso. Si se conoce una solución x_2, ..., x_n de estas $m-1$ ecuaciones últimas de (E'), se hallará una solución para el sistema completo, pues la primera ecuación permite expresar x_1 en función de las demás incógnitas, como

$$x_1=b_1'-c_{12}x_2-c_{13}x_3-\ldots-c_{1n}x_n$$

Por inducción relativa al número m de ecuaciones, esto demuestra:

TEOREMA 9. *Todas las soluciones posibles de un sistema de ecuaciones lineales simultáneas con coeficientes de un campo, pueden hallarse con un número finito de operaciones racionales (adiciones, sustracciones, multiplicaciones y divisiones) en ese campo.*

¿Cuál es el resultado final de las reducciones sucesivas de (E) a (E')? En la reducción k-ésima, una de las variables no eliminadas antes, llamémosla x_k, será eliminada, y el problema se reduce a un sistema de $m-k$ ecuaciones con $n-k$ incógnitas. Este proceso sólo terminará cuando no queden ecuaciones, o cuando sean nulos los coeficientes de las incógnitas de las ecuaciones que quedan. Así pues, el sistema toma la forma

$$(\mathrm{E}') \qquad \begin{aligned} x_1 + d_{12}x_2 + d_{13}x_3 + \dots + d_{1n}x_n &= b_1' \\ x_2 + d_{23}x_3 + \dots + d_{2n}x_n &= b_2' \\ x_3 + \dots + d_{3n}x_n &= b_3' \\ \dots\dots\dots\dots\dots\dots & \\ x_r + \dots + d_{rn}x_n &= b_r' \qquad (r \leqslant m) \end{aligned}$$

en el cual $r=m$ o $r<m$, y en este último caso las restantes ecuaciones tienen la forma $0=b_i'$; entonces, estas ecuaciones últimas serán absurdas (no tendrán solución) a menos que todas las constantes b_i' sean cero para $i>r$. En el sistema (E'), todas las soluciones se hallan dando valores arbitrarios a las variables $x_{r+1}, \dots, x_n$, resolviendo la última ecuación para x_r, la penúltima para x_{r-1}, y así sucesivamente, hasta obtener x_1. En el caso particular de ser $m=n=r$, el sistema (E') tiene forma triangular, y hay una solución única.

Un modo práctico de reducir las ecuaciones (E) a (E') es omitir las variables y escribir sólo los coeficientes en forma de cuadro ordenado:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{bmatrix} \rightarrow \begin{bmatrix} 1 & d_{12} & d_{13} & \dots & d_{1n} & b_1' \\ & 1 & d_{23} & \dots & d_{2n} & b_2' \\ & & \dots & \dots & \dots & \dots \end{bmatrix}$$

Las operaciones que ordinariamente se aplican a las ecuaciones pueden igualmente aplicarse al cuadro correspondiente. Por ejemplo, para restar tres veces la primera *ecuación* de la cuarta, se resta simplemente tres veces la primera *fila* de la cuarta. Estos cuadros, llamados matrices, se estudiarán más adelante con mayor detalle (Capítulos VIII y X).

Un sistema de ecuaciones (E) es *homogéneo* si las constantes b_i de los segundos miembros son todas nulas. Tales sistemas tienen siempre la solución trivial $x_1 = x_2 = \ldots = x_n = 0$. Puede ocurrir que no tengan más soluciones, pero si el número de variables excede al de ecuaciones, la última ecuación de (E') contendrá siempre alguna variable «extra», que podrá tomar un valor arbitrario. Además, las ecuaciones absurdas $0 = b_i'$ no pueden nunca aparecer en las ecuaciones homogéneas. Por lo tanto,

TEOREMA 10. *Un sistema de* m *ecuaciones lineales homogéneas con* n *variables y* $m < n$ *tiene siempre alguna solución en que no son nulas todas las variables.*

EJERCICIOS

1. Resolver las siguientes congruencias simultáneas:

   a) $3x+2y \equiv 1$ (mód. 7); $4x+6y \equiv 3$ (mód. 7).

   b) $2x+7y \equiv 3$ (mód. 11); $3x+4z \equiv 6$ (mód. 11); $4x+7y+z \equiv 0$ (mód. 11).

   c) $x-2y+z \equiv 5$ (mód. 13); $2x+2y \equiv 7$ (mód. 13);
   $5x-3y+4z \equiv 1$ (mód. 13).

2. Resolver las ecuaciones a) y b), con el módulo suprimido, en el campo R de los números racionales.

3. Resolver en $R(\sqrt{2})$ las ecuaciones simultáneas

$$(1+\sqrt{2})x+(1-\sqrt{2})y=2; \qquad (2-\sqrt{2})x+(3-\sqrt{2})y=1$$

4. Hallar todas las soluciones incongruentes de las congruencias simultáneas

$$x+y+z \equiv 0 \text{ (mód. 5)}; \qquad 3x+2y+4z \equiv 0 \text{ (mód. 5)}$$

5. Hallar todas las soluciones incongruentes de las congruencias simultáneas

   a) $x+2y-z+5t \equiv 4$; $2x+5y+z+2t \equiv 1$; $x+3y+2z+6t \equiv 2$
   (todas mód. 7).

   b) $x+y+z \equiv 1$ (mód. 5); $3x+3y+3z \equiv 4$ (mód. 5).

6. Demostrar que dos ecuaciones $a_1x_1+\ldots+a_nx_n=c$, $b_1x_1+\ldots+b_nx_n=d$, tienen siempre solución para coeficientes en un campo dado, supuesto que no existan constantes $k \neq 0$ y $m \neq 0$ tales, que $ka_i = mb_i$ para $i=1, \ldots, n$.

7. Demostrar que si $(x_1, \ldots, x_n)$ es una solución de un sistema homogéneo de ecuaciones lineales, también $(-x_1, \ldots, -x_n)$ será solución. ¿Qué puede decirse de la suma de dos soluciones?

4. Campos ordenados

Se dice que un campo está ordenado, si contiene un conjunto de elementos «positivos» con las propiedades aditiva, multiplicativa y tricotómica expuestas en el § 3 del Capítulo I; en otras palabras, un campo está ordenado cuando, considerado como un dominio, es un dominio de integridad ordenado. Sabemos por aritmética que los números racionales constituyen un campo ordenado, pero vamos ahora a probarlo partiendo de nuestra construcción de los racionales como pares de enteros; después mostraremos que el método «natural» de ordenación es el único modo de formar con los números racionales un campo ordenado

Recordemos que en cualquier dominio ordenado al cuadrado b^2 de un elemento no nulo, es siempre positivo. Si un cociente a/b es positivo, el producto $(a/b)b^2 = ab$ debe ser también positivo y recíprocamente. De aquí que en un campo ordenado,

(11) $a/b > 0$ si, y sólo si, $ab > 0$.

Pero dijimos que el número racional (a, b) representaba el cociente a/b. Así pues, diremos que un número (a, b) es *positivo* si el producto ab es positivo, y sólo en este caso.

TEOREMA 11. *Los números racionales forman un campo ordenado si* $(a, b) > 0$ *significa, por definición, que el entero* ab *es positivo.*

Demostración. Como hemos definido la igualdad por un convenio, hemos de demostrar que $(a, b) > 0$ y $(a, b) \equiv (c, d)$ implican $(c, d) > 0$. Esto es cierto, puesto que cd tiene el mismo signo que b^2cd, ab el mismo signo que abd^2, y $abd^2 = b^2cd$, en virtud de la hipótesis $ad = bc$. Entre los elementos positivos se verifican también las necesarias propiedades referentes a la adición, multiplicación y tricotomía (Capítulo I, § 3). Por ejemplo, la suma de dos pares positivos, (a, b) y (c, d), es positiva, pues $ab > 0$ y $cd > 0$ implican $d^2ab > 0$ y $b^2cd > 0$, de donde deducimos

$$bd(ad + bc) = d^2ab + b^2cd > 0,$$

lo que quiere decir que la suma $(ad+bc, bd)$ es positiva. Finalmente, la definición de fracciones «positivas» conserva el orden natural en las fracciones particulares $(a, 1)$, que representan enteros, ya que $(a, 1)$ es positivo, por la definición (11), sólo cuando $1\cdot a > 0$.

Como en la demostración del Teorema 11 la única propiedad de los enteros que se utiliza es la de que forman un dominio ordenado, podremos enunciar el siguiente resultado general :

TEOREMA 12. *El campo* F *de cocientes de un dominio de integridad* D *puede ser ordenado, conviniendo en que el cociente* a/b *de elementos de* D *es positivo si, y sólo si,* ab *es positivo. Este es el único modo de extender el orden de* D *a un orden de* F.

Existen otros campos ordenados : el campo de los números reales, el campo $R(\sqrt{2})$ de números $a+b\sqrt{2}$ (ver § 1) y otros subcampos del campo de los números reales. En cualquiera de tales campos puede introducirse un concepto de valor absoluto como en § 3, Capítulo I, y todas las propiedades de las desigualdades se conservan. Además de las reglas válidas en todo dominio de integridad, se puede demostrar :

(12) $$0 < 1/a \quad \text{cuando} \quad a > 0,$$

(13) $$a/b < c/d \quad \text{si, y sólo si,} \quad abd^2 < b^2cd,$$

(14) $$0 < a < b \quad \text{implica} \quad 0 < 1/b < 1/a,$$

(15) $$a < b < 0 \quad \text{implica} \quad 0 > 1/a > 1/b,$$

(16) $$a_1^2 + a_2^2 + \dots + a_n^2 \geqslant 0.$$

Las dos reglas (14) y (15) son las usuales para la división de desigualdades. La regla (16) dice que una suma de cuadrados no es nunca negativa (Teorema 2, Capítulo I) y es utilizada con frecuencia. Por ejemplo, si $a \neq b$, entonces $(a-b)^2 > 0$, así que $a^2 - 2ab + b^2 > 0$, lo cual da $a^2 + b^2 > 2ab$. Poniendo ahora $x = a^2$, $y = b^2$, y dividiendo por 2, queda :

$$(x+y)/2 > \sqrt{xy} \qquad (x \neq y).$$

Esto nos dice que la media aritmética de dos números positivos es mayor que su media geométrica.

También puede probarse el recíproco del Teorema 11 :

TEOREMA 13. *Todo campo ordenado* F *contiene un subcampo que es isomorfo con el campo* R *de los números racionales.*

Por el Teorema 20 del Cap. I, F contiene un subdominio isomorfo con el dominio de los enteros. El subcampo engendrado por este dominio será isomorfo con R, por el Teorema 6. Finalmente, por la fórmula (11) y el hecho de que cada suma $n = 1^2 + 1^2 + \ldots + 1^2$ debe ser positiva [cfr. (16)] y cada $-(1^2 + 1^2 + \ldots + 1^2)$ negativa, vemos que el concepto de positivo tiene en nuestro subcampo el significado usual.

Este resultado da una caracterización abstracta de campo de los números racionales como *el menor campo ordenado*.

EJERCICIOS

1. Demostrar que en cualquier campo ordenado $ab > 0$ implica $a/b > 0$.
2. Admitiendo que los enteros formen un dominio ordenado, demostrar que el producto de dos números racionales positivos es positivo.
3. De modo semejante, demostrar que si $(a, b) \not\equiv 0$, una de las alternativas $(a, b) > 0$ y $-(a, b) > 0$ será válida.
4. Demostrar $|xx' + yy'| \leqslant \sqrt{(x^2+y^2)(x'^2+y'^2)}$ en cualquier campo ordenado en el que todos los números positivos tengan raíces cuadradas. (*Sugerencia:* Cuadrar ambos miembros de la igualdad.)
5. Demostrar las fórmulas (12)-(16) del texto.
6. Si n es entero positivo, y a y b son números racionales positivos, demostrar que $(a^n + b^n)/2 \geqslant [(a+b)/2]^n$. (*Sugerencia:* poner $(a+b)/2 = r$, $a = r + d$, $b = r - d$.)
7. a) Demostrar que cualquier subcampo de un campo ordenado es un campo ordenado.
   b) ¿Es cualquier subdominio de un campo ordenado un dominio ordenado?
8. Para los números racionales (o, más generalmente, en cualquier campo ordenado) demostrar que, si $a < b$, existirán infinitos x satisfaciendo a $a < x < b$.
9. Demostrar que en un campo desordenado, los elementos positivos forman un conjunto bien ordenado.

* 5. Axiomática del número natural

Hemos utilizado el sistema J de los enteros como punto de partida hacia otros sistemas numéricos fundamentales. Podríamos haber utilizado para lo mismo el sistema más restringido J^+ de los enteros positivos (números naturales).

De los postulados que hemos dado para caracterizar el sistema J de los enteros, se deducen algunas propiedades de los números naturales, que bastan para caracterizarlos axiomáticamente.

TEOREMA 14. *El sistema* J^+ *constituido por todos los enteros positivos de* J, *tiene las siguientes propiedades:*

1.ª Es cerrado respecto a dos operaciones binarias, adición y multiplicación, definidas univocamente y que cumplen las leyes asociativa, conmutativa y distributiva.

2.ª Existe en J^+ *un elemento 1, idéntico para la multiplicación, tal que* $m \cdot 1 = m$ *para cualquier* m *de* J^+.

3.ª Son válidas en J^+ *las siguientes leyes de simplificación:*

$$\text{(17)} \qquad \textit{Si es } m + x = n + x, \quad \textit{debe ser } \quad m = n.$$

$$\text{(18)} \qquad \textit{Si es } m \cdot x = n \cdot x, \quad \textit{debe ser } \quad m = n.$$

4.ª Si m *y* n *son dos números cualesquiera de* J^+, *vale una, y sólo una, de las siguientes alternativas:* $m = n$, o $m + x = n$ *tiene una solución* x *en* J^+, *o* $m = n + y$ *tiene una solución* y *en* J^+.

5.ª Finalmente, es válido en J^+ *el principio de inducción completa: cualquier subconjunto de* J^+ *que contenga al 1, y que si contiene al* n *contenga también al* $n + 1$, *contendrá a todos los elementos de* J^+.

Dejamos la demostración de estas propiedades de J^+ como ejercicio para el lector.

Recíprocamente, si tomamos como postulados las propiedades 1.ª a 5.ª ahora enunciadas, quedan caracterizados precisamente los números naturales, es decir: los enteros positivos, según antes los hemos definido, cumplen estas propiedades, y cualquier otro sistema de entes que cumplan estos postulados coincide, salvo isomorfismos, con el de los enteros positivos. Obsérvese, en particular, que la ley de simplificación para la suma se ha tomado ahora como postulado, ya que no es posible deducirla de la posibilidad de la sustracción, como se hizo en el § 2 del Capítulo I. Asimismo, las tres alternativas sobre las ecuaciones $m + x = n$, ocupan el lugar de alguna de las propiedades de orden de los enteros positivos.

Partiendo de los números naturales, como definidos por los anteriores postulados, se puede reconstruir el sistema de los enteros.

El objeto de esta construcción es hallar un sistema más amplio que el J^{+} en el cual sea siempre posible la sustracción. Para esto, introduciremos como elementos nuevos, ciertos pares (m, n) de enteros positivos, cada uno de los cuales representa la solución de la ecuación respectiva $n+x=m$. Los detalles de este proceder son semejantes a los expuestos en la construcción de los racionales a partir de los enteros (§ 2).

DEFINICIÓN. *Un entero es un par* (m, n) *de números naturales* m *y* n. *La igualdad de dos pares se define por el convenio*

$$(19) \qquad (\mathrm{m}, \mathrm{n}) \equiv (\mathrm{r}, \mathrm{s}) \quad \textit{significa que} \quad \mathrm{m}+\mathrm{s}=\mathrm{n}+\mathrm{r},$$

y las sumas y productos se definen por

$$(20) \qquad (\mathrm{m}, \mathrm{n})+(\mathrm{r}, \mathrm{s})=(\mathrm{m}+\mathrm{r}, \mathrm{n}+\mathrm{s}),$$

$$(21) \qquad (\mathrm{m}, \mathrm{n}) \cdot (\mathrm{r}, \mathrm{s})=(\mathrm{mr}+\mathrm{ns}, \mathrm{ms}+\mathrm{nr}).$$

Por último, (m, n) *es positivo si, y sólo si,* n + x = m, *para algún número natural* x.

Los pares introducidos con esta definición satisfacen, en efecto, todos los postulados que dimos para los enteros. Debemos observar primero que la igualdad introducida por (19) es reflexiva, simétrica y transitiva, y que la suma y producto dados por (20 y (21) están definidos unívocamente respecto a esta relación de igualdad (consultar los razonamientos del § 11 del Capítulo I). Las varias leyes formales de los dominios de integridad resultan válidas, y ello se ve por aplicación sistemática de las definiciones (20) y (21), de modo semejante a lo que se hizo para los números racionales. En particular, (2, 1) es una unidad, y (1, 1), un cero, del sistema ahora definido. El aditivo inverso existe, pues

$$(m, n)+(n, m) \equiv (1, 1) \quad \text{para cualquier } (m, n).$$

Vemos así que estos pares forman un dominio de integridad.

Por el postulado 4.° del Teorema 14, cualquier par puede escribirse precisamente en una de las tres formas (n, n), $(n+x, n)$ o $(n, n+x)$. Los de la primera forma son iguales al cero (1, 1); los de la segunda forma $(n+x, n)$ son los pares positivos, y puede demostrarse que verifican las propiedades aditiva, multiplicativa y de tricotomía, requeridas en la definición de un dominio de integri-

dad (Capítulo I, § 3). Además, $(m+x,\ m) \equiv (n+y,\ n)$ si, y sólo si, $x=y$. Por lo tanto, si identificamos los pares «congruentes», la correspondencia $x \leftrightarrow (n+x,\ n)$ es biunívoca, entre los números naturales y los pares positivos $(n+x,\ n)$. Se trata, precisamente, de un isomorfismo, puesto que por la definición (20)-(21),

$$(m+x,\ m)+(n+y,\ n)=(m+n+x+y,\ m+n),$$

$$(m+x,\ m)\cdot(n+y,\ n)=$$
$$=(mn+my+nx+mn+xy,\ mn+nx+mn+my).$$

Por lo tanto, los pares «positivos» satisfacen al principio de inducción completa. Con esto queda bosquejada la demostración del siguiente resultado :

TEOREMA 15. *El sistema* J^+ *de los enteros positivos puede sumergirse en un sistema más amplio* J, *en el que es posible la sustracción, y de tal modo, que cualquier elemento de* J *es la diferencia de dos enteros positivos de* J^+. *El sistema* J *así construído es un dominio de integridad ordenado, cuyos elementos positivos satisfacen la ley de inducción completa.*

Por Capítulo I, § 5, Ejercicio 8, este resultado implica el principio de buena ordenación. Puede notarse que en la demostración aquí apuntada intervienen sólo los postulados para J^+. Recíprocamente, en cualquier dominio que contenga al J^+, las diferencias $(a-b)$ de los elementos de J^+ deben satisfacer a las definiciones (19)-(21). (Cfr. Capítulo I, § 2, Ejercicio 3.) Esto demuestra

TEOREMA 16. *Cualquier dominio de integridad que contenga al sistema* J^+, *debe contener un subdominio isomorfo con el dominio* J *de todos los enteros.*

Para una construcción sistemática de la aritmética a partir de los números naturales, resulta deseable un sistema de postulados más simple ; no como el enunciado ahora, mediante dos operaciones binarias, adición y multiplicación, sino basado en una sola operación *unitaria* $S(n)=n+1$. Esta función $S(n)$ se lee *sucesivo de* n. Los postulados de Peano, relativos a esta función, son :

1.°, 1 es un número. 2.°, Cualquier número n determina un número único $S(n)$ como sucesor. 3.°, Ningún número tiene al 1

como sucesor. 4.°, $S(m)=S(n)$ implica $m=n$. 5.°, El principio de inducción completa.

Partiendo de los «números naturales» así definidos, la adición y la multiplicación se introducen por definiciones de la forma

$$(22) \qquad n+1=S(n), \qquad n+S(m)=S(n+m);$$

$$(23) \qquad n\cdot 1=n, \qquad n\cdot S(m)=n\cdot m+n.$$

Estas definiciones son *recurrentes*, es decir, hacen posible calcular un producto $n(m+1)$ mediante otro producto conocido, cuyo segundo factor es más pequeño.

EJERCICIOS

1. Demostrar que la relación definida por (19) es reflexiva, simétrica y transitiva.
2. Demostrar que si $(m,n)\equiv(m',n')$, también $(m,n)+(r,s)\equiv(m',n')+(r,s)$ y $(m,n)\cdot(r,s)\equiv(m',n')\cdot(r,s)$, para todo (r,s).
3. Demostrar que la «adición» definida por (20) es conmutativa y asociativa.
4. Demostrar lo mismo para la «multiplicación» definida por (21).
5. Demostrar que (m, m) es la misma para todo m, y es el cero aditivo. ¿Es la primera proposición consecuencia de la segunda?
6. Demostrar que $(m+1, m)$ es elemento idéntico para la multiplicación.
7. Demostrar la ley distributiva.
8. Demostrar la ley de simplificación para la multiplicación.
9. ¿Qué propiedades de $J+$ se han utilizado en los ejercicios 1-8? Enunciar un teorema que tenga la misma relación con los Teoremas 15-16 que el Teorema 7 con los Teoremas 5-6.
10. Justificar la definición de «positivo» para el par (m, n).
11. Demostrar detalladamente el Teorema 14.
12. Demostrar que el postulado IV del Teorema 14 puede reemplazarse por la condición de que $m+1\neq 1$ para cualquier m de $J+$. (Este es, en realidad, el tercer postulado de Peano.)
13. Definir en $J+$, $m<n$ cuando $m+x=n$ para algún $x\in J+$. Probar: a) $m<n$ y $n<r$ implica $m<r$; b) $m<m$ para ningún m; c) $m<n$ implica $m+r<n+r$ para todo r; d) $m<n$ implica $mr<nr$ para todo r.

*14. Demostrar que las condiciones c) y d) del ejercicio 13 pueden utilizarse para reemplazar las leyes de simplificación (17)-(18) en la lista de postulados para $J+$.

*15. Demostrar que los postulados de Peano valen para los elementos positivos en cualquier dominio ordenado en el que los elementos positivos estén bien ordenados.

*16. a) Dados los postulados de Peano y la definición 22, demostrar por inducción que $n+1=1+n$.
b) Demostrar que la adición definida por (22) es conmutativa.

*17. Establecer un conjunto de postulados para el sistema R^+ de todos los *números racionales positivos.* (*Sugerencia:* Hallar primero los postulados para los elementos positivos en cualquier campo ordenado, y luego caracterizar a R^+ como el menor de estos sistemas.)

18. Demostrar que la repetición del proceso utilizado para obtener J a partir de J^+ no vale para nuevas extensiones de J. ¿Puede generalizarse este resultado?

CAPITULO III

Números reales

1. Dilema de Pitágoras

El filósofo griego Pitágoras (hacia el 600 a. de C.) sabía ya que la razón $r = d/l$ entre la longitud d de la diagonal de un cuadrado y la longitud l de su lado, satisface a la igualdad

(1) $$d^2 = (rl)^2 = l^2 + l^2$$ (Teorema de Pitágoras)

Así pues, razonaba él, existe un «número» r tal, que $r^2 = 1 + 1 = 2$.

Pero, por otra parte, Pitágoras reconoció que r no podía representarse como un cociente $r = a/b$ de enteros. En efecto, $(a/b)^2 = 2$ implicaría $a^2 = 2b^2$. Mas descomponiendo a en factores primos, resulta que a^2 es divisible por 2 un número par de veces; y por lo análogo, 2 dividirá a $2b^2$ un número impar de veces. Luego $a^2 = 2b^2$ es imposible para a y b enteros.

Únicamente podemos solucionar este «dilema de Pitágoras» introduciendo los números *irracionales*: números que no son cocientes de enteros.

Razonamientos similares (que el lector podrá suplir) demuestran que la razón $\sqrt{3}$ entre la longitud de la diagonal de un cubo C y la longitud de su arista, y la razón $\sqrt[3]{2}$ entre la longitud de una arista de C y la del cubo de volumen mitad, son números irracionales. Estos resultados son casos particulares del siguiente teorema mucho más general:

TEOREMA 1. *Sea* $\mathrm{p}(\mathrm{x}) = \mathrm{x}^n + \mathrm{a}_1\mathrm{x}^{n-1} + \ldots + \mathrm{a}_n$ *un polinomio con su primer coeficiente igual a 1 y los demás* $\mathrm{a}_1, \ldots, \mathrm{a}_n$ *enteros. Si la ecuación* $\mathrm{p}(\mathrm{x}) = 0$ *tiene raíces racionales, éstas son enteras.*

Supongamos que $p(x)=0$ para alguna fracción $x=a/b$. Dividiendo a y b por su m. c. d. puede expresarse x como un cociente $x=r/l$ de dos enteros r, l, primos entre sí. Sustituyendo este valor en $p(x)$ y quitando denominadores,

$$(2) \qquad 0=l^n p(r/l)=r^n+a_1 r^{n-1}l+a_2 r^{n-2}l^2+\ldots+a_n l^n$$

luego $r^n=-a_1 r^{n-1}l-\ldots-a_n l^n$, luego $l \mid r^n$. Esto exige que cualquier factor primo de l divida a r^n y por tanto a r. Pero r y l no tienen divisores comunes, y por tanto $l=\pm 1$ y la fracción dada $x=r/(\pm 1)=\pm r$ es un número entero, c. q. d.

Para probar la irracionalidad de $\sqrt{28}$, por ejemplo, fundándonos en el Teorema 1, procederemos como sigue: si $|x|\geqslant 6$, $x^2-28>0$; si $|x|\leqslant 5$, $x^2-28<0$; luego ningún entero puede ser solución de $x^2-28=0$ y, por el Teorema 1, la solución $\sqrt{28}$ de $x^2=28$ no puede ser racional.

Otros números irracionales son π (que no es, por tanto, exactamente igual a 22/7 ni a 3'1416), e y muchos otros. En el Cap. XIV probaremos no sólo que la inmensa mayoría de los números reales son irracionales, sino incluso que, a diferencia de $\sqrt{2}$, no pueden satisfacer a ninguna ecuación algebraica. Este resultado, que hemos anticipado, nos indica ya que para contestar a la pregunta: ¿qué es un número real?, necesitaremos utilizar ideas enteramente nuevas.

La naturaleza de estas ideas y la relación entre los números reales y los racionales serán examinadas parcialmente en los párrafos que siguen.

EJERCICIOS

1. Demostrar: si un polinomio $p(x)=x^n+a_1x^{n-1}+\ldots+a_n$ con coeficientes enteros, tiene una raíz entera r, este r es divisor de a_n.
2. Demostrar que $\sqrt{3}$ es irracional sin utilizar el Teorema 1.
3. Demostrar que la ecuación $2x^3+x=5$ no tiene raíces racionales. (*Sugerencia:* Considerar la nueva variable $y=2x$.)
4. Comprobar si las siguientes ecuaciones tienen raíces racionales:

   a) $3x^3-7x=5$; b) $5x^3+x^2+x=4$;

   c) $8x^3+3x^2=17$; d) $6x^3-3x=18$.
5. Demostrar que $20x^n=91$ no tiene ninguna raíz racional para $n>1$ (entero). (*Sugerencia:* Utilizar el teorema del factor primo.)
6. * ¿Para qué números racionales x es $3x^2-7x$ entero? Hallar condiciones necesarias y suficientes.

7. ¿Para qué enteros a entre 0 y 250 tiene raíces racionales la ecuación $30x^n=a$ para algún $n>1$?

8. Demostrar que $\sqrt[n]{a}$ es irracional excepto si el entero a es la n-ésima potencia de algún entero.

9. Demostrar que $\log_{10} 3$ es irracional. (*Sugerencia:* Recordar la definición de logaritmo.)

*10. Demostrar que $\sqrt{2}+\sqrt{5}$ es irracional. (*Sugerencia:* Hallar la ecuación polinómica correspondiente.)

*11. a) Si u es irracional, mientras que a y b son racionales, ¿cuándo será racional $au+b$?

b) Si también c y d son racionales, ¿cuándo será racional

$$(au+b)/(cu+d)?$$

2. Números reales. Método geométrico y expresión decimal

Los griegos de la época clásica usaron un método geométrico de aproximación para el cálculo de los números reales. Para ellos, un número era simplemente una razón $(a:b)$ entre dos segmentos rectilíneos a y b. En consecuencia, dieron construcciones geométricas para establecer la igualdad entre razones, así como para la adición, sustracción, multiplicación y división de razones. De este modo, las leyes del Algebra (Cap. II, § 1) aparecen como teoremas geométricos.

La versión griega de la noción de igualdad entre números racionales y reales se basaba en una condición, debida a Eudoxio, que especificaba cuándo eran iguales dos razones (*). Esta condición se hacía depender de las posibilidades de formar geométricamente los múltiplos enteros $m \cdot a$ de un segmento dado a, y de comparar geométricamente las longitudes de dos segmentos. Se estipulaba que $(a:b)=(c:d)$ cuando, para todo par de enteros positivos m y n,

(3) Si $ma>nb$, también $mc>nd$; si $ma<nb$, también $mc<nd$

Algebraicamente, $ma>nb$ significa que $a/b>n/m$, suponiendo siempre que b y m sean positivos. Entonces (3) puede leerse así: $a/b=c/d$, cuando cualquier número racional n/m que sea menor que a/b es también menor que c/d, mientras cualquier n/m que sea mayor que a/b es también mayor que c/d.

(*) El lector podrá hacerse idea del papel que desempeña esta condición si recuerda las primeras proposiciones sobre semejanza en la geometría euclídea.

Supongamos que $p(x)=0$ para alguna fracción $x=a/b$. Dividiendo a y b por su m. c. d. puede expresarse x como un cociente $x=r/l$ de dos enteros r, l, primos entre sí. Sustituyendo este valor en $p(x)$ y quitando denominadores,

(2) $$0=l^n p(r/l)=r^n+a_1 r^{n-1} l+a_2 r^{n-2} l^2+\ldots+a_n l^n$$

luego $r^n=-a_1 r^{n-1} l-\ldots-a_n l^n$, luego $l \mid r^n$. Esto exige que cualquier factor primo de l divida a r^n y por tanto a r. Pero r y l no tienen divisores comunes, y por tanto $l=\pm 1$ y la fracción dada $x=r/(\pm 1)=\pm r$ es un número entero, c. q. d.

Para probar la irracionalidad de $\sqrt{28}$, por ejemplo, fundándonos en el Teorema 1, procederemos como sigue: si $|x| \geqslant 6$, $x^2-28>0$; si $|x| \leqslant 5$, $x^2-28<0$; luego ningún entero puede ser solución de $x^2-28=0$ y, por el Teorema 1, la solución $\sqrt{28}$ de $x^2=28$ no puede ser racional.

Otros números irracionales son π (que no es, por tanto, exactamente igual a 22/7 ni a 3'1416), e y muchos otros. En el Cap. XIV probaremos no sólo que la inmensa mayoría de los números reales son irracionales, sino incluso que, a diferencia de $\sqrt{2}$, no pueden satisfacer a ninguna ecuación algebraica. Este resultado, que hemos anticipado, nos indica ya que para contestar a la pregunta: ¿qué es un número real?, necesitaremos utilizar ideas enteramente nuevas.

La naturaleza de estas ideas y la relación entre los números reales y los racionales serán examinadas parcialmente en los párrafos que siguen.

EJERCICIOS

1. Demostrar: si un polinomio $p(x)=x^n+a_1 x^{n-1}+\ldots+a_n$ con coeficientes enteros, tiene una raíz entera r, este r es divisor de a_n.
2. Demostrar que $\sqrt{3}$ es irracional sin utilizar el Teorema 1.
3. Demostrar que la ecuación $2x^2+x=5$ no tiene raíces racionales. (*Sugerencia:* Considerar la nueva variable $y=2x$.)
4. Comprobar si las siguientes ecuaciones tienen raíces racionales:

   a) $3x^3-7x=5$; b) $5x^3+x^2+x=4$;

   c) $8x^3+3x^2=17$; d) $6x^3-3x=18$.
5. Demostrar que $20x^n=91$ no tiene ninguna raíz racional para $n>1$ (entero). (*Sugerencia:* Utilizar el teorema del factor primo.)

* 6. ¿Para qué números racionales x es $3x^2-7x$ entero? Hallar condiciones necesarias y suficientes.

7. ¿Para qué enteros a entre 0 y 250 tiene raíces racionales la ecuación $30x^n = a$ para algún $n > 1$?

8. Demostrar que $\sqrt[n]{a}$ es irracional excepto si el entero a es la n-ésima potencia de algún entero.

9. Demostrar que $\log_{10} 3$ es irracional. (*Sugerencia:* Recordar la definición de logaritmo.)

10. Demostrar que $\sqrt{2}+\sqrt{5}$ es irracional. (*Sugerencia:* Hallar la ecuación polinómica correspondiente.)

11. a) Si u es irracional, mientras que a y b son racionales, ¿cuándo será racional $au+b$?
    b) Si también c y d son racionales, ¿cuándo será racional

$$(au+b)/(cu+d)?$$

2. Números reales. Método geométrico y expresión decimal

Los griegos de la época clásica usaron un método geométrico de aproximación para el cálculo de los números reales. Para ellos, un número era simplemente una razón $(a : b)$ entre dos segmentos rectilíneos a y b. En consecuencia, dieron construcciones geométricas para establecer la igualdad entre razones, así como para la adición, sustracción, multiplicación y división de razones. De este modo, las leyes del Algebra (Cap. II, § 1) aparecen como teoremas geométricos.

La versión griega de la noción de igualdad entre números racionales y reales se basaba en una condición, debida a Eudoxio, que especificaba cuándo eran iguales dos razones (*). Esta condición se hacía depender de las posibilidades de formar geométricamente los múltiplos enteros $m \cdot a$ de un segmento dado a, y de comparar geométricamente las longitudes de dos segmentos. Se estipulaba que $(a : b) = (c : d)$ cuando, para todo par de enteros positivos m y n,

(3) Si $ma > nb$, también $mc > nd$; si $ma < nb$, también $mc < nd$

Algebraicamente, $ma > nb$ significa que $a/b > n/m$, suponiendo siempre que b y m sean positivos. Entonces (3) puede leerse así: $a/b = c/d$, cuando cualquier número racional n/m que sea menor que a/b es también menor que c/d, mientras cualquier n/m que sea mayor que a/b es también mayor que c/d.

(*) El lector podrá hacerse idea del papel que desempeña esta condición si recuerda las primeras proposiciones sobre semejanza en la geometría euclídea.

La validez de la condición (3) de Eudoxio expresa, evidentemente, la circunstancia de que dos números reales positivos $(a:b)$ y $(c:d)$ son diferentes si, y sólo si, existe algún número racional mayor que uno de ellos y menor que el otro. También su condición para $(a:b)<(c:d)$ tiene el mismo fundamento, y es la siguiente:

(4) $ra<lb$ y $rc>ld$, para enteros convenientes r y l.

El estudio geométrico de los números reales es ya desacostumbrado. En la actualidad se les estudia aritméticamente, mediante aproximaciones racionales, en especial decimales (un decimal es, como se sabe, un número racional cuyo denominador es potencia de 10). Por ejemplo, el irracional $\sqrt{2}$ se reemplaza en la práctica por las aproximaciones sucesivas 1, 1'4, 1'41, 1'414, ... El número π es aproximado, análogamente, por los decimales

(5) $d_1=3'1$, $d_2=3'14$, $d_3=3'141$, $d_4=3'1415$, $d_5=3'14159$,

y así sucesivamente.

Al número π se le llama extremo superior del conjunto infinito de números racionales (5), porque tal número es, por lo menos, tan grande como cualquiera de los términos de la sucesión (5), mientras que ningún número real menor que π tiene esta propiedad. La existencia de «extremos superiores» es una propiedad característica de los números reales, como vamos a explicar en el § 3.

EJERCICIOS

1. Demostrar que $x=0,12437437437...$ representa un número racional. (*Sugerencia:* Calcular $1000x-x$.)
2. Lo mismo para $y=1,23672367...$
3. Demostrar que cualquier decimal periódico, semejante à los de los ejercicios 1 y 2, representa un número racional. Calcular su numerador y denominador.
4. Demostrar que, inversamente, el desarrollo decimal de cualquier número es periódico. (*Sugerencia:* Observar que si aparece el mismo resto después de m y de $m-k$ divisiones por 10, el grupo de k cifras entre ambos se repite indefinidamente.)

*5. ¿Es válido el resultado del ejercicio 4 en la base duodecimal?

6. Hallar tres aproximaciones sucesivas de $\sqrt{2}$ en el dominio de todos los números racionales con denominadores potencias de 3.

3. Postulados de los números reales

Mediante el concepto de extremo superior, y con postulados convenientes, podremos introducir ahora los números reales.

DEFINICIÓN. *Diremos que un campo ordenado* F *es completo cuando todo subconjunto* S *de* F *no vacío y acotado superiormente tiene un extremo superior en* F.

Aclaración Se llama «cota superior» de S a cualquier elemento de F no excedido por ningún elemento de S, esto es, a cualquier elemento b tal que $x < b$ para todo x de S. Una cota superior b se llama extremo superior si no hay ningún elemento menor que b que sea cota superior de S; esto es, si para cualquier $b' < b$ existe en S un x tal que $b' < x$. Así, extremo superior es sinónimo de «cota superior mínima» (brevemente, c. s. m.). De modo análogo se definen la «cota inferior» y el extremo inferior o cota inferior máxima (c. i. m.).

Postulado de los números reales. Los números reales forman un campo R^* ordenado y completo.

En este postulado quedan incluídas todas las propiedades de los números reales, ya que pueden deducirse de él. Así sucede, por ejemplo, con el Teorema de Rolle, fundamental para probar el Teorema de Taylor y otras cuestiones de Cálculo Infinitesimal. En la próxima sección daremos, como ejemplo, algunas demostraciones de sencillas propiedades de los números reales, en las que interviene la existencia de tales extremos superiores. En el § 5 demostraremos que dos campos cualesquiera ordenados y completos son isomorfos, así que la afirmación de que los números reales R^* constituyen un campo completo y ordenado, es una caracterización abstracta de R^*, exactamente lo mismo que, como vimos, los postulados para los enteros los caracterizaron de modo abstracto (Capítulo I, § 13). Pero ahora nos limitaremos a mostrar más explícitamente el porqué se hace intervenir a los extremos superiores. Es posible definir concretamente el campo R^* de los números reales como un conjunto de desarrollos decimales de infinitas cifras (a los que, por brevedad, llamaremos *decimales ilimitados*) com-

binados y ordenados según reglas apropiadas (*). Pero si se hace esto, es muy difícil probar lo que se supone implícitamente en el álgebra de la escuela media, que el sistema R^* de estos decimales ilimitados es un *campo ordenado* (¡confróntense nuestros postulados para campos ordenados con las hipótesis del álgebra que hemos estudiado en la escuela!). Por otra parte, es relativamente fácil demostrar que un conjunto S, no vacío, de decimales ilimitados, acotado superiormente, tiene un extremo *superior* a^* decimal ilimitado.

En primer lugar, cualquier conjunto T no vacío, de decimales ilimitados, no negativos, tiene un extremo inferior c. Para mostrarlo consideremos los números con n cifras decimales que resultan limitando en sus n primeras cifras decimales los números de T. Entre ellos habrá uno mínimo, ya que los números (no negativos) con n cifras decimales y menores que cualquier número de T forman conjunto finito. Sea este mínimo $k+0'd_1d_2...d_n$, en el que k es un entero y $d_1, ..., d_n$ son dígitos. El mínimo de los números análogos con $n+1$ cifras coincide con el antedicho en sus n primeras cifras, así que tiene la forma $k+0'd_1d_2...d_nd_{n+1}$, con un nuevo dígito d_{n+1}. Nuestra construcción define así cierto número decimal ilimitado $k+0'd_1d_2...$ Por la construcción resulta ser una *cota inferior* de T (ya que por su desarrollo decimal no es mayor que ningún x de T), y además es un *extremo inferior* (pues con cualquier cifra decimal mayor en su desarrollo no tendría aquella propiedad).

Podemos, por lo tanto, hallar un extremo inferior para todo conjunto no vacío S acotado inferiormente. Sea b una cota inferior; si restamos b de cada número y de S, obtendremos un conjunto T de números no negativos $y-b$, y para este conjunto puede hallarse un extremo inferior c, como anteriormente. El número $b^*=c+b$ es entonces un extremo inferior para el conjunto original S, como puede verse fácilmente.

Con este resultado pueden construirse también extremos superiores. Si un conjunto S tiene una cota superior a, el conjunto de los opuestos de los elementos de S tendrá la cota inferior $-a$;

(*) Las primeras complicaciones se deben a igualdades como 0'199...=0'200... entre decimales distintos. También resulta difícil dar una regla para deducir con un número finito de pasos la n-ésima cifra de una suma o producto. Esto se refleja en las complicaciones cuando se quiere demostrar que el sistema así definido es un campo (por ejemplo, probar que vale la ley asociativa o que existe el inverso $1/x$).

luego, por la demostración anterior, tendrá un extremo inferior b^* El número $a^* = -b^*$ es un extremo superior de S. Vemos así que los decimales ilimitados constituyen un sistema «completo», en el sentido de nuestra definición.

En la axiomática de los números reales pudimos también haber postulado la existencia de extremos inferiores. La única razón para no haberlo hecho, es que esto puede demostrarse partiendo de las restantes hipótesis.

TEOREMA 2. *En el campo* R^* *de los números reales (definido por nuestros postulados como un campo ordenado y completo) todo subconjunto* S *no vacío que tiene una cota inferior, tiene un extremo inferior.*

La correspondencia $x \to (-x)$ es una correspondencia biunívoca de R^* consigo mismo; y como transforma $x < y$ en $-x > -y$, cambiará también cotas superiores en inferiores y extremos superiores en inferiores. El postulado de existencia de extremos superiores encierra, pues, la de los inferiores.

EJERCICIOS

1. Sea D un dominio ordenado en el que cualquier conjunto acotado tiene una c. s. m.; demostrar que también tiene c. i. m. (un conjunto está «acotado» si tiene cota superior y cota inferior).
2. Demostrar que el principio de buena ordenación implica la existencia de c. s. m. y c. i. m. en los conjuntos acotados de enteros.
3. Demostrar que no hay dominio ordenado en el que *cualquier* conjunto no vacío tenga una c. s. m. (*Sugerencia:* Demostrar que el propio D no puede tener cota superior.)
4. Cualquier conjunto *finito* no vacío de números racionales tiene c. s. m. y c. i. m. ¿Es cierto esto para cualquier dominio ordenado?
5. Establecer en lenguaje geométrico un postulado sobre los puntos del eje real, que asegure que todo conjunto acotado tiene c. s. m. y c. i. m. (Utilizar las palabras «izquierda» y «derecha».)
6. ¿Cuál es la c. s. m. del conjunto de todos los números racionales m/n, con $m^2 < 7n^2$?
7. Escribir dos conjuntos racionales diferentes que tengan ambos la misma c. s. m. $\sqrt{2}$.
8. Mostrar la c. s. m. de cada uno de los siguientes conjuntos de números racionales:

a) 1/3, 4/9, 13/27, 40/81, ...; b) 1/2, 3/4, 7/8, 15/16, ...

9. Supongamos que S tiene una c. s. m. a^* y una c. i. m. b^*.
   a) Demostrar detalladamente que el conjunto de todos los números $-3x$, con x en S, tiene la c. s. m. $-3b^*$ y la c. i. m. $-3a^*$.
   b) Del mismo modo, hallar la c. s. m. y la c. i. m. del conjunto de todos los números $x+5$ para x en S.
10. En el Ejercicio 9 calcular la c. s. m.: a) del conjunto de todos los números $7x+2$, con x en S; b) del conjunto de todos los números $1/x$, con $x \neq 0$ en S, si $b^* > 0$.
11. Sean S_1 y S_2 dos conjuntos de números reales con las respectivas c. s. m. b_1 y b_2. ¿Cuál es la c. s. m.: a) del conjunto S_1+S_2 de todas las sumas s_1+s_2, con $s_1 \in S_1$ y $s_2 \in S_2$; b) del conjunto de todos los elementos que pertenecen a S_1 o a S_2?
12. Escribir una relación de todos los postulados para los números reales.
*13. Construir un sistema de postulado para los números reales positivos. (*Sugerencia:* Cfr. Cap. 2, § 5.)
14. Mostrar que cualquier elemento a^* en un campo ordenado es c. s. m. para un conjunto S si, y sólo si, 1) $x \leqslant a^*$ para todo $x \in S$; 2) para cualquier e positivo del campo, existe un x en S con $|x-a^*| < e$.

4. Aplicación de los postulados

Vamos ahora a utilizar la existencia de extremos superiores para demostrar varias propiedades del campo de los números reales R^*, comenzando por la existencia de soluciones para ecuaciones tales como $x^2=2$.

TEOREMA 3. *Si* p(x) *es un polinomio con coeficientes reales, si* a < b *y si* p(a) < p(b), *entonces, para cada constante* C *que satisfaga a la limitación* p(a) < C < p(b), *la ecuación* p(x) = C *tiene al menos una raíz entre* a *y* b.

Geométricamente, la hipótesis significa que la gráfica de $y=p(x)$ une los puntos $[a, p(a)]$ y $[b, p(b)]$; la conclusión asegura que dicha gráfica debe cortar toda horizontal $y=C$, intermedia entre $y=p(a)$ e $y=p(b)$ en algún punto con coordenada x entre a y b.

Demostración. Consideremos en primer lugar el caso especial $p(x)=x^2$, con $a=1$, $b=2$ y $C=2$. Como $a^2=1<C<4=b^2$, la conclusión nos afirma que la ecuación $x^2=2$ tiene una solución real entre 1 y 2. En efecto, indiquemos con S el conjunto de los números reales comprendidos entre 1 y 2 que satisfacen a la condición $x^2 \leqslant 2$. Como $1^2=1 \leqslant 2$, S no es vacío; y S es acotado, pues todos sus elementos son menores que 2; tendrá, pues, un extremo superior c. Vamos a demostrar que $c^2=2$.

Observemos primero que si Δx es un número real (un incremento de x),

$$(6) \qquad (c+\Delta x)^2 = c^2 + 2c\Delta x + (\Delta x)^2 = c^2 + (2c+\Delta x)\Delta x.$$

Pero si $|\Delta x| \leqslant 1$, resultará $|2c+\Delta x| \leqslant 2|c|+1$. Si $M = 2|c|+1$, tendremos

$$(7) \qquad |(c+\Delta x)^2 - c^2|/M \leqslant |\Delta x|, \quad \text{si} \quad |\Delta x| \leqslant 1.$$

(Esta fórmula nos dice, esencialmente, que el incremento $(c+\Delta x)^2 - c^2$ nunca excede en M veces al correspondiente incremento de x.)

Veamos ahora que $c^2 > 2$ es imposible. Pues en tal caso, cualquier elemento x de S podría escribirse $x = c - \Delta x$, con Δx no negativo, así que $x^2 = (c - \Delta x)^2 \leqslant 2$, por la definición de S. Pero por (7), $\Delta x = |\Delta x| \geqslant |(c - \Delta x)^2 - c^2|/M = [c^2 - (c - \Delta x)^2]/M \geqslant (c^2 - 2)/M$. Por consiguiente, todos los elementos x de S entre 1 y 2 satisfarán

$$x = c - \Delta x \leqslant c - (c^2 - 2)/M,$$

así que $c - (c^2 - 2)/M$ sería una cota superior de S, contrariamente a la hipótesis de que c es el extremo superior.

Análogamente, $c^2 < 2$ es imposible. Pues ello implicaría por (7), con $\Delta x = (2 - c^2)/M$, que

$$[c + (2 - c^2)/M]^2 \leqslant c^2 + M(2 - c^2)/M = 2,$$

contrariamente a la construcción de c como cota superior del conjunto de todos los x, con $x^2 \leqslant 2$.

El razonamiento en el caso general es idéntico, excepto esto: en las fórmulas (6) y (7) se tiene $p(c+\Delta x) = p(c) + q(c, \Delta x)\Delta x$, con un polinomio más general $q(c, \Delta x)$, y además, para obtener M, se debe poner, en $q(c, \Delta x)$, $|b - a|$ en vez de Δx, $|c|$ en vez de c, y signo más en todos los términos.

COROLARIO 1. *Si* p(x) *es un polinomio con coeficientes positivos y sin término independiente, y además* C > 0, *la ecuación* p(x) = C *tendrá al menos una raíz positiva.*

Como no hay término independiente, $p(0) = 0$, mientras que $0 < C < p(b)$ para un valor de b suficientemente grande.

COROLARIO 2. *Si* p(x) *es de grado impar, entonces* p(x)=C *tiene al menos una raíz para cualquier valor real del número* C.

Pues si $p(x)$ tiene positivo su primer coeficiente, se verifica que $p(-b) < C < p(b)$ para valores suficientemente grandes de b. Si fuese negativo el coeficiente del primer término, resolveríamos $-p(x) = -C$ como en el primer caso.

Nota. El Teorema 3 no enseña el modo de calcular numéricamente las soluciones de las ecuaciones numéricas; se limita a probar su existencia. Para calcularlas puede emplearse la regla de Newton, como se enseña en los cursos de Cálculo, o los desarrollos en serie como el

$$(8) \qquad \sqrt{1+x} = 1 + \frac{1}{2}x + \frac{1}{2}\left(-\frac{1}{2}\right)\frac{x^2}{2!} + \frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\frac{x^3}{3!} + \dots$$

Este estudio no corresponde al Álgebra, sino al Análisis.

Nuestros postulados hacen del conjunto de los números reales un campo ordenado R^*, y el Teorema 13 del Capítulo II muestra que R^* debe contener un subcampo isomorfo con el campo R de los racionales. Como R fué definido en el Cap. II a menos de isomorfismos, podemos suponer que el campo R^* de los números reales contiene a todos los racionales, y por ende a todos los enteros. Este convenio ajusta nuestros postulados en la aritmética de uso corriente y nos permite probar la siguiente propiedad de los números reales (llamada comúnmente «propiedad arquimedeana»):

TEOREMA 4. *Dados dos números cualesquiera,* $a > 0$ *y* $b > 0$, *en el campo* R^* *de los números reales (definido por los precedentes postulados), existe un entero* n *para el cual* $na > b$.

Demostración. Supongamos la conclusión falsa para un par de números reales a y b, así que $b \geqslant na$ para cualquier n. El conjunto S de todos los múltiplos na es acotado (b es una cota), así que tendrá un extremo superior b^*. Entonces $b^* \geqslant na$ para cualquier n, así que también $b^* \geqslant (m+1)a$ para cualquier m. Esto implica $b^* - a \geqslant ma$, luego $b^* - a$ es una cota superior para el conjunto S de todos los múltiplos de a, a pesar de ser menor que su extremo superior b^*, lo cual es contradictorio.

El establecimiento de la propiedad arquimedeana puede utilizarse para demostrar la condición de Eudoxio [cfr. § 2, (4)].

TEOREMA 5. *Entre dos números reales cualesquiera* $c > d$ *existe un número racional* m/n *tal que* $c > m/n > d$.

Como antes, lo probaremos basándonos en que las números reales forman un campo ordenado y completo (postulado). Por hipótesis, $c - d > 0$, así que por la propiedad arquimedeana existirá un entero positivo n tal que $n(c-d) > 1$, o bien, $1/n < c - d$. Sea ahora m el menor entero para el cual $m > nd$; entonces $(m-1)/n \leqslant d$; así que

$$m/n = (m-1)/n + 1/n < d + (c-d) = c.$$

Y como $m/n > d$, resulta $c > m/n > d$, como queríamos demostrar.

Podemos ilustrar esta demostración como sigue: las diversas fracciones 0, $\pm 1/h$, $\pm 2/h$, ..., con denominador fijo h, están distribuídas a lo largo del eje real con intervalos de longitud $1/h$. Para estar seguros de que uno de tales puntos cae entre c y d necesitamos sólo hacer que el espacio $1/h$ sea menor que la diferencia $c - d$.

Este teorema puede servirnos para sustentar rigurosamente nuestra idea intuitiva, utilizada, por ejemplo, en las aproximaciones (5), de que todo número real es extremo superior de un conjunto de números racionales.

COROLARIO. *Todo número real es extremo superior de un conjunto de números racionales.*

Demostración. Para un número real dado c, sea S el conjunto de todos los números racionales $m/n \leqslant c$. En tal caso, c es una cota superior de S. Por el teorema, ningún número real $d < c$ puede serlo, luego c es extremo superior de S, c. q. d.

EJERCICIOS

1. Demostrar que cualquier número positivo real tiene una raíz cuadrada real.
2. Probar que cualquier número real tiene raíz cúbica real.
3. Demostrar que $x^4 - x = C$ tiene dos raíces reales para todo $C > -3/8$.
*4. Demostrar que si $\operatorname{sen} a < \operatorname{sen} b$, para todo C que satisfaga la limitación $\operatorname{sen} a < C < \operatorname{sen} b$, la ecuación $\operatorname{sen} x = C$ tiene una raíz real entre a y b. (*Sugerencia:* Obsérvese que $|\operatorname{sen}(x+\Delta x) - \operatorname{sen} x| \leqslant |\Delta x|$.)
5. Calcular $\sqrt{5}$ con cuatro decimales, utilizando (8) y observando que $(1/2\sqrt{5})^2 = 1 + 1/4$.
6. Calcular $\sqrt{2}$ con seis cifras, utilizando (8) y $(5\sqrt{2}/7)^2 = 1 + 1/49$.

7. Calcular $\sqrt{17}$ con cinco decimales.
8. Demostrar que entre dos números reales cualesquiera $c < d$, existe un cubo racional $(m/n)^3$ tal que $c < (m/n)^3 < d$. ¿Es verdad esto para los cuadrados racionales?
9. Si $h > 1$ es entero, demostrar que entre dos números reales $c > d$ existe un número racional de la forma m/h^k, con m y k enteros.
10. Demostrar que la igualdad $a/b = c/d$ entre dos cocientes de números reales se define correctamente pór la condición de Eudoxio dada en (3), § 2.
11. a) Si a y b son positivos reales, mostrar que $ax^{n+1} > bx^n$ para todo valor positivo suficientemente grande de x.
    b) Dado un polinomio $p(x)$ con coeficiente principal positivo, hallar un número real M tal que $p(x) > 0$ para todo $x > M$.
12. Traducir el argumento del Teorema 3 al caso $x^3 - 7$, $a = 1$, $b = 2$.

*13. Desarrollar el razonamiento del Teorema 3 en caso general.

* 5. Cortaduras de Dedekind

Imaginemos a los números racionales distribuídos en la forma habitual sobre el eje de las x. Si cortamos este eje (como con unas tijeras), dividiremos a los números racionales en dos clases : llamemos L a la clase de la izquierda y U a la de la derecha. Cualquier número racional se encuentra en una de las dos clases, y puede suceder que *uno* de estos números, r/s, sea el borde de ambas. Esto sucederá si, y sólo si, el eje de las x se ha cortado precisamente por el punto r/s. Obsérvese especialmente que si x está en L, será $x \leqslant y$ para cualquier y de U. Viceversa, si $x \leqslant y$ para todos los y de U, x estará en L. Tenemos con esto una idea de lo que es una *cortadura de Dedekind*.

Para precisar esta idea, sea F un campo ordenado. Se llama *cortadura* de Dedekind en F, a un par de subconjuntos no vacíos L y U tales, que

1.° L es el conjunto de todas las cotas inferiores de los elementos de U;

2.° U es el conjunto de todas las cotas superiores de los elementos de L. (Cfr. el próximo Ejercicio 5.)

LEMA 1. *En las partes inferior y superior de una cortadura de Dedekind, tomadas en conjunto, quedan incluídos todos los elementos de* F ; *ambas tienen, a lo más, un elemento común.*

Demostración. Sea dado un $x \in F$. Si $x \leqslant a$, para algún $a \in L$, será $x \leqslant a \leqslant y$ para todo y de U, y, por lo tanto, $x \in L$. En otro

caso, por la ley tricotómica, $x > a$ para toda $a \in L$, así que $x \in U$; esto demuestra la primera afirmación: cualquier elemento de F está en L o en U. Además, sean a y b dos elementos que están contenidos simultáneamente en L y en U. Será $a \geqslant b$ (puesto que $a \in U$, $b \in L$), y $a \leqslant b$ (puesto que $a \in L$, $b \in U$), y, por lo tanto, debe ser $a = b$, lo que demuestra la segunda afirmación.

Cuando L y U tengan el elemento a en común diremos que la cortadura se ha hecho *a través* de a. Evidentemente, una cortadura (L_a, U_a) estará hecha a través de a, cuando L_a sea el conjunto de los $x \leqslant a$, y U_a el conjunto de los $x \geqslant a$.

Axioma de Dedekind (*para cortaduras sobre un campo ordenado* F). *Toda cortadura está hecha a través de algún elemento* a.

Teorema 6. *El axioma de Dedekind es válido en un campo ordenado* F *si, y sólo si, se trata de un campo ordenado y completo* (*).

Demostración. Sea (L, U) una cortadura. Supongamos primero que F es completo. Por definición (§ 3), la clase L contiene a su extremo superior a. Como a es una cota superior de L (precisamente, la cota superior *mínima*), pertenecerá a U; pero por ser la *mínima*, resulta una cota inferior de las cotas superiores, es decir, de los elementos de U. Esto significa (por la definición de cortadura) que a pertenece también a L; por consiguiente, la cortadura (L, U) ha sido hecha a través de a.

Recíprocamente, supongamos válido el axioma de Dedekind, y sea S un conjunto no vacío y acotado. Llamemos U al conjunto de todas las cotas superiores de S, y L al conjunto de todas las cotas inferiores de U (por lo tanto, L contiene a S). Para probar que (L, U) es una cortadura, bastará mostrar que U es, exactamente, el conjunto de las cotas superiores de L. Pero, por la construcción de L, cualquier elemento de U es una cota superior de L ($x \leqslant y$ para todo $x \in L$, $y \in U$); además, U incluye a *todas* estas cotas superiores, puesto que L contiene a S. Probado ya que (L, U) es una cortadura, el axioma de Dedekind nos dice que estará hecha a través de un elemento a. Este a, por ser elemento de U, será cota

(*) Que F sea un campo no es esencial. MacNeille demostró en 1935 que el más general conjunto parcialmente ordenado puede ser completado mediante las cortaduras de Dedekind, para formar una «lattice» completa. (Cfr. Cap. XI.)

superior de S, y por ser elemento de L, será la cota superior mínima ($a \leqslant x$ para todo $x \in U$). Esto completa la demostración.

Podemos ahora apuntar una demostración sobre la naturaleza característica del postulado enunciado en § 3 para los números reales.

TEOREMA 7. *Dos campos ordenados y completos son isomorfos.*

Demostración. Sean tales campos F' y F''. Por el Teorema 13 del Cap. II, contendrán sendos subcampos «racionales», R' y R'', isomorfos (isomorfismo que conserva el orden, así como las sumas y productos). Extenderemos el isomorfismo entre R' y R'' a un isomorfismo entre F' y F''.

En efecto, cualquier $a' \in F'$ define una cortadura en F' y, mediante ella, una cortadura en R' (esto es, en el subcampo de los racionales). Pero por el Teorema 5 (condición de Eudoxio), a' está determinado por esta cortadura en R', y asimismo, cualquier cortadura (L_R, U_R) en R' determina un elemento a' que es extremo superior para L_R y extremo inferior para U_R. Las cortaduras en R'' se definen análogamente, y de este modo, los elementos de F' y los de F'' se corresponden biunívocamente con las respectivas cortaduras en R' y en R'', y por ende, biunívocamente entre sí.

Finalmente, las operaciones en F' y en F'' pueden definirse mediante las de R' y R'', de modo que se extienda el isomorfismo. Precisando más, sean a y b los elementos correspondientes a las cortaduras (L_a, U_a) y (L_b, U_b) de R'. En tal caso, $a+b$ corresponderá a la cortadura (L_a+L_b, U_a+U_b) (*) —designando por L_a+L_b el conjunto de las sumas $x+y$, con $x \in L_a$ e $y \in L_b$, y lo análogo para U_a+U_b—. Para multiplicar elementos *positivos*, a y b, formemos de modo análogo las respectivas cortaduras en el sistema de los racionales positivos. Entonces ab corresponde a la cortadura (L_aL_b, U_aU_b) —donde L_aL_b es el conjunto de los productos xy, con $x \in L_a$ e $y \in L_b$, y lo análogo para U_aU_b. Como $(-a)b = a(-b) = -ab$, y $(-a)(-b) = ab$, la extensión alcanza a todos los productos. Omitimos los detalles de la demostración.

Recíprocamente, se pueden utilizar las cortaduras para definir

(*) En ciertos casos (L_a+L_b, U_a+U_b) no será una cortadura de R, porque el número $a+b$ sea racional y, sin embargo, no aparezca en ninguna de las dos clases; pero la cortadura aparece sin más que añadir a ambas partes el número desaparecido. Lo mismo se dice para el producto, que sigue ahora.

los números reales a partir de los enteros. De los postulados del dominio J de los enteros se deduce (cfr. Cap. II, § 5) que los racionales forman un campo ordenado con la propiedad arquimedeana enunciada en Teorema 4. Si se definen la adición y la multiplicación de cortaduras en R del modo expuesto en los precedentes párrafos, se puede demostrar que las cortaduras en R constituyen un campo ordenado que satisface el axioma de Dedekind, y que es, por lo tanto, un campo completo y ordenado. Pero las demostraciones son largas, y su desarrollo nos llevaría lejos de nuestro camino, así que nos limitamos a enunciar el resultado.

TEOREMA 8. *Existe un campo completo ordenado, y sólo uno, salvo isomorfismos.*

EJERCICIOS

1. Demostrar que si (L, U) y (L′, U′) son cortaduras en el campo racional, cualquier número racional con una excepción a lo más, puede escribirse como $x+y$ ($x \in$ L, $y \in$ L′) o como $u+v$ ($u \in$ U, $v \in$ U).
2. Enunciar y demostrar teoremas análogos para los números racionales *positivos*, con la multiplicación en vez de la adición.
3. ¿Por qué fallan estos teoremas para los negativos racionales?
4. Demostrar que para todo $\epsilon > 0$ existe un n bastante grande para que $10^{-n} < \epsilon$.
5. A veces se define una cortadura de Dedekind en un campo ordenado F como un par de subconjuntos L′ y U′ de F tales, que cualquier elemento de F esté siempre en L′ o en U′, y tal que $x < y$ siempre que $x \in$ L′ e $y \in$ U′. Por adición y supresión de convenientes números particulares, demostrar que cualquier cortadura (L′, U′) de este tipo da una cortadura (L, U) en sentido del texto, y viceversa.
6. Si t es un elemento de un dominio ordenado D con $0 < t < 1$, demostrar que $s = 2 - t$ tienen las propiedades $s > 1$, $st \leqslant 1$.
7. Sea D un dominio ordenado «completo». a) Si D no es isomorfo con J, mostrar que D contiene un elemento t con $0 < t < 1$. b) Si b y c son elementos positivos cualesquiera de D, demostrar que t^nb $<$ c para algún n.

8. Mediante los Ejercicios 6 y 7 demostrar que cualquier dominio ordenado «completo» es isomorfo o bien a J o bien a R. (*Sugerencia:* Hallar el inverso de b > 1 considerando todos los x con xb $\leqslant 1$.)

9. a) Demostrar que cualquier isomorfismo de R* consigo mismo conserva la relación $x \leqslant y$. (*Sugerencia:* $x \leqslant y$ si, y sólo si, $z^2 = y - x$ tiene una raíz.)
   b) Utilizando a), demostrar que el único isomorfismo de R* consigo mismo es el isomorfismo trivial $x \to x$.

* 6. Convergencia

Los números reales tienen otra interpretación, distinta de la que acabamos de exponer mediante las cortaduras. Se les puede considerar, en efecto, como *límites* de *sucesiones* de números racionales. Este punto de vista se asocia al nombre de G. Cantor, y se basa en la siguiente

DEFINICIÓN. *Se dice que una sucesión* $\{x_n\} = x_1, x_2, x_3, \ldots$, *de elementos de un campo ordenado, converge hacia un límite* a *en* F (*en símbolos,* $x_n \to a$), *cuando, dado cualquier* $\epsilon > 0$ *en* F, *existe un* $N = N(\epsilon)$ *tal, que para todo* $n > N$ *sea* $|x_n - a| < \epsilon$.

Dicho de otro modo, todos los términos de la sucesión, excepto un número finito de ellos, satisfacen la desigualdad $|x_n - a| < \epsilon$. Por ejemplo, la sucesión 3, 3'1, 3'14, 3'141, ..., de § 2, (5), converge hacia π, porque el n-ésimo término de esta sucesión difiere de π en menos de $0'00...01 = 10^{-n}$. En este ejemplo, podemos tomar concretamente $N = N(\epsilon)$ mayor que $-\log \epsilon$; con lo cual, todos los términos x_n, con $n > N$, difieren de π en menos de ϵ. Este ejemplo ilustra el sentido con que cada número real aparece como límite de una sucesión de racionales (que en la práctica son decimales).

LEMA 1. *Si* $x_n \to a$, *existe un entero* N *tal, que* $|x_m - x_n| < \epsilon$ *siempre que sean* $n > N$ *y* $m > N$.

Demostración. Tomemos N lo bastante grande para que $k > N$ implique $|x_k - a| < \epsilon/2$. Entonces, si $m > N$ y $n > N$, resultará $|x_m - x_n| = |(x_m - a) + (x_n - a)| \leqslant |x_m - a| + |x_n - a| < \epsilon$. La condición del enunciado se debe a Cauchy, y su importancia radica en que en ella no interviene a. Llamaremos *sucesión regular* (o *sucesión de Cauchy*) a toda sucesión $\{x_n\} = x_1, x_2, x_3, \ldots$, tal, que $|x_m - x_n| \to 0$ cuando $m, n \to \infty$ en la forma que expresa el Lema.

TEOREMA 9. *Cualquier sucesión regular de números reales (o bien de elementos de un campo ordenado y completo) tiene límite.*

Demostración. Sea $\{x_n\}$ la sucesión regular. Ante todo, x_n permanece acotado; pues si tomamos $\epsilon = 1$ y $N = N(1)$, los términos que siguen a x_N tienen como cota superior $x_N + 1$, luego la totalidad de la sucesión admitirá como cota superior el mayor de los

números $x_1, x_2, \ldots, x_{N-1}, x_N+1$. Lo mismo se establece que hay una cota inferior. Podemos, pues, formar para cada n el número $y_n =$ extremo superior de $(x_n, x_{n+1}, x_{n+2}, \ldots)$; se demuestra que $y_1 \geqslant y_2 \geqslant y_3 \geqslant \ldots$ está acotado por las mismas cotas que $\{x_n\}$. Sea a el extremo inferior de las y_n. Vamos a demostrar que $x_n \to a$.

En efecto, sea un $\epsilon > 0$ dado. Por definición de a, $a \leqslant y_N \leqslant a+\epsilon/3$, y como $\{x_n\}$ es una sucesión regular, si N es suficientemente grande, para todo m, $n \geqslant N$ será $|x_m - x_n| < \epsilon/3$. Además, por la definición de y_N, podemos tomar $n \geqslant N$ tal, que $y_N \geqslant x_n \geqslant y_N - \epsilon/3$; combinando con lo anterior, para todo $m > N$, resulta:

(8) $$|x_m - a| \leqslant |x_m - x_n| + |x_n - y_N| + |y_N - a| < \epsilon,$$

como queríamos demostrar.

El Teorema 9 ocupa en la construcción de los números reales un papel semejante al axioma de Dedekind (*). Tiene además una importancia posterior, que no comparte el último: se aplica también a los números complejos (Cap. V), a los números p-ádicos (que no tratamos en este libro), a la geometría y al análisis.

La construcción de Cantor de los números reales, a partir de los racionales, se basa en las siguientes definiciones:

1.ª Toda sucesión regular de números racionales «es» un número real;

1.ª' Dos sucesiones regulares $\{x_n\}$ e $\{y_n\}$ son «equivalentes» cuando $|x_n - y_n| \to 0$;

2.ª $\{x_n\} + \{y_n\} = \{x_n + y_n\}$;

3.ª $\{x_n\}\{y_n\} = \{x_n y_n\}$.

La construcción de Cantor es menos algebraica y más artificiosa que la de Dedekind. (Por ejemplo, como en la construcción de los racionales en el Capítulo II, se debe demostrar que la adición y la multiplicación son unívocas respecto a la equivalencia introducida.) Sin embargo, la demostración de que el sistema definido por 1.ª - 3.ª es un campo, resulta mucho más breve (**) que la correspondiente demostración con las cortaduras de Dedekind, mientras que, por otra parte, se relaciona de modo más estrecho con la teoría de las series infinitas y otros métodos generales del Análisis.

(*) Es fácil demostrar que un campo arquimedeano satisface al axioma de Dedekind si, y sólo si, todas sus sucesiones regulares tienen límite.

(**) Puede verse el Capítulo VI de C. C. Mac Duffee, *Introduction to Abstract Algebra*. Wiley, 1940.

EJERCICIOS

1. Demostrar que si $\{x_n\}$ e $\{y_n\}$ son sucesiones regulares, también lo son a) $\{x_n+y_n\}$, b) $\{x_ny_n\}$, c) $\{-x_n\}$, d) $\{1/y_n\}$, supuesto en la última que y_n no converja a 0, y que ningún $y_n=0$.
2. Sean $\{x_n\}\equiv\{y_n\}$ definidas como en 1.ª' en el texto. Demostrar que esta relación es reflexiva, simétrica y transitiva.
3. En Ejerc. 2, demostrar que $\{x_n\}\equiv\{y_n\}$ implica $\{x_n\}+\{a_n\}\equiv\{y_n\}+\{a_n\}$ y $\{x_n\}\cdot\{a_n\}\equiv\{y_n\}\cdot\{a_n\}$ para toda $\{a_n\}$ regular.
4. Demostrar las leyes conmutativa, asociativa y distributiva y la existencia de aditivos inversos para sucesiones regulares.
5. Hacer lo mismo para el inverso multiplicativo de una sucesión $\{x_n\}\equiv 0$. (*Nota:* Puede aparecer un número finito de términos $x_n=0$.)

* 6. Para un primo fijo p, sea $\| a \|$ el valor absoluto p-ádico del entero a, definido en el Capítulo I, § 8, Ejercicio 3.
   a) Definir la convergencia y regularidad de las sucesiones de enteros, utilizando siempre $\| x \|$ en lugar de $| x |$.
   b) Demostrar el análogo del Lema 1.
   c) Demostrar que si la igualdad, suma y producto se definen convenientemente, las sucesiones regulares forman un dominio de integridad (al que se le llama dominio de los *enteros p-ádicos*).

* 7. Demostrar que cualquier entero p-ádico puede expresarse unívocamente como una serie

$$a_0+a_1p+a_2p^2+\dots,\qquad 0\leqslant a_i<p$$

y mostrar cómo se calculan las sumas y productos de tales series.

CAPITULO IV

Polinomios

Formas polinómicas

Sea D un dominio de integridad, y sea «x» simplemente un mbolo. Supongamos que se forman sumas, productos y diferenıs de x con los elementos de D y consigo mismo, con sujeción las reglas ordinarias del cálculo algebraico (es decir, de acuerdo n los postulados que caracterizan a un dominio de integridad). n casos particulares, este procedimiento es conocido por el Álgea de la Enseñanza Media, y nos lleva a la construcción de *poliımios en x*. Mientras no se haga ninguna hipótesis sobre x (ni n siquiera la de que sea un elemento incógnito de D), se suele cir que x es una *indeterminada*.

Evidentemente, el procedimiento descrito nos permite construir das las expresiones de la forma

$$) \qquad a_0+a_1x+\ldots+a_nx^n \qquad (a_0,\ \ldots,\ a_n \in D\ ;\ a_n \neq 0 \text{ si } n>0)$$

ı las que x^n (n entero positivo) se define como $x \cdot x \ldots x$ (n factos). Y asimismo, utilizando sólo los postulados del dominio de tegridad, se pueden sumar, restar o multiplicar dos expresiones ıalesquiera de la forma (1), obteniendo otra expresión de la misa forma. Por ejemplo, si D es el dominio de los enteros,

$$\begin{aligned} x) &= [0+1\cdot x+(-2)x^2](2+3\cdot x)= \\ &=0\cdot 2+0\cdot 3\cdot x+1\cdot x\cdot 2+1\cdot x\cdot 3\cdot x+(-2)x^2\cdot 2+(-2)x^2\cdot 3\cdot x= \\ &=0+0\cdot x+2x+3x^2+(-4)x^2+(-6)x^3= \\ &=0+(0+2)x+[3+(-4)]x^2+(-6)x^3= \\ &=0+2x+(-1)x^2+(-6)x^3, \end{aligned}$$

por la ley distributiva generalizada, las leyes conmutativa y asociativa y, finalmente, la ley distributiva.

Este razonamiento puede generalizarse. En efecto, sean

$$p(x) = a_0 + a_1 x + \dots + a_m x^m$$

y

$$q(x) = b_0 + b_1 x + \dots + b_n x^n$$

dos expresiones de la forma (1). Si $m \geqslant n$, tendremos

(2) $$p(x) \pm q(x) = (a_0 \pm b_0) + \dots + (a_n \pm b_n)x^n + a_{n+1}x^{n+1} + \dots + a_m x^m.$$

Resulta una fórmula análoga si $m < n$. Además, por la ley distributiva, $p(x)q(x) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_i b_j x^{i+j}$. Agrupando los términos con el mismo exponente, y sumando los coeficientes, tenemos

(3) $$p(x)q(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \dots + a_m b_n x^{m+n}$$

En este resultado, el coeficiente de x^k es una suma $\sum_i a_i b_{k-i}$ extendida a los valores enteros de i tales, que $0 \leqslant i \leqslant m$, y $0 \leqslant k - i \leqslant n$ (fig. 1). Con esto hemos demostrado el siguiente resultado:

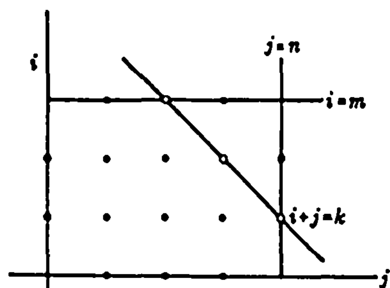


Figura 1

TEOREMA 1. *Sea D' un dominio de integridad (supuesto existente) tal, que contenga a un subdominio isomorfo con otro dominio dado D y a un elemento x no perteneciente a D. Entonces los polinomios (1) en este elemento x se suman, restan y multiplican por las fórmulas (2) y (3), y constituyen un subdominio de D'.*

Con objeto de demostrar que siempre existe un tal dominio de integridad D', daremos la siguiente definición:

DEFINICIÓN. *Llamaremos forma polinómica (o polinomio formal) en una indeterminada x sobre un dominio de integridad D, a toda expresión del tipo (1). Se llama grado de la forma (1) al entero n. Diremos que dos polinomios (formales) son iguales cuando son del mismo grado y tienen iguales los coeficientes de las mis-*

mas potencias de x. *Polinomio* (*) *nulo es el que tiene iguales a cero todos sus coeficientes.*

Teorema 2. *Si la adición y la multiplicación se definen por las fórmulas (2) y (3), las diferentes formas polinómicas en* x *sobre un dominio de integridad* D, *formarán un nuevo dominio de integridad* D[x] *que contiene a* D.

La ausencia de divisores de cero (ley de simplificación en la multiplicación) se deduce de (3), ya que el coeficiente $a_m b_n$ del primer término del producto de dos polinomios no nulos, es el producto de los primeros coeficientes a_m y b_n de los factores, y por ende no será nulo. Las propiedades del 0 y del 1 y la existencia de opuestos, se deducen fácilmente de (2) y (3).

Para demostrar las leyes conmutativa, asociativa y distributiva es conveniente introducir «falsos» coeficientes nulos. Esto da para (2) y (3) las expresiones más simples

$$(2') \qquad \sum_{k=0}^{\infty} a_k x^k + \sum_{k=0}^{\infty} b_k x^k = \sum_{k=0}^{\infty} (a_k + b_k) x^k,$$

$$(3') \qquad \left(\sum_{k=0}^{\infty} a_k x^k\right)\left(\sum_{k=0}^{\infty} b_k x^k\right) = \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i b_j x^k\right),$$

donde sólo un número finito de coeficientes no son nulos. Cualquier ley formal, como la distributiva, por ejemplo, puede ser comprobada fácilmente operando según (2')-(3') en ambos miembros de su expresión, o sea

$$\left(\sum_k a_k x^k\right)\left(\sum_k b_k x^k + \sum_k c_k x^k\right) = \sum_k \left[\sum_{i+j=k} a_i(b_j + c_j)\right] x^k$$

$$\left(\sum_k a_k x^k\right)\left(\sum_k b_k x^k\right) + \left(\sum_k a_k x^k\right)\left(\sum_k c_k x^k\right) = \left(\sum_k \left[\left(\sum_{i+j=k} a_i b_j\right) + \left(\sum_{i+j=k} a_i c_j\right)\right] x^k.$$

Por ser válida en *D* la ley distributiva, el coeficiente de la *k*-ésima potencia de x es el mismo en ambas expresiones. Análogos razonamientos completan la prueba del Teorema 2.

Recordando ahora el Teorema 7 del Capítulo II, vemos que si

(*) Aquí, y en lo sucesivo, cuando no haya lugar a confusión, diremos simplemente «polinomio», sobreentendiendo el calificativo «formal». — N. del T.

definimos una *forma racional* en la indeterminada x sobre D, como un *cociente* formal

$$\frac{p(x)}{q(x)}=\frac{a_0+a_1x+\ldots+a_mx^m}{b_0+b_1x+\ldots+b_nx^n}\quad (a_i,\ b_j \text{ en } D\ ;\ a_m\neq 0 \text{ si } m>0,\ b_n\neq 0)$$

de *formas polinómicas* con denominador no nulo, y definimos la igualdad, adición y multiplicación, según (5), (6) y (7) de la definición del § 2, Capítulo II, obtenemos un campo.

COROLARIO. *Las formas racionales en una indeterminada* x *sobre cualquier dominio de integridad* D, *constituyen un campo. Este campo se denota por* D(x).

EJERCICIOS

1. Reducir a la forma (1): $x^3-5x(3x+7)^2$, $(x^3+5x-4)(x^2-2x+3)$, $(3x^3+7x-1/2)(x^3-x/2+1)$.
2. Calcular de modo semejante $(3x^3+5x-4)(4x^3-x+3)$, donde los coeficientes son enteros mód. 7.
3. ¿Es x^3+5x-4 de la forma (1)? Reducirlo a esta forma. Reducir $(1+x+2x^2+3x^3)-(0+x+x^3+3x^2)$ a la forma (1), destacando qué postulados se usan en cada paso.
4. a) ¿Es $1/2+3\cdot x^{1/2}+5x$ una forma polinómica sobre el campo racional?
   b) ¿Por qué $x^3\cdot x^4$ no es igual a x^2 en el dominio de las formas polinómicas con coeficientes en J_5?
5. Discutir las siguientes proposiciones:
   a) El grado del producto de dos formas polinómicas es la suma de los grados de los factores.
   b) El grado de la suma de dos formas polinómicas es el mayor grado de los sumandos.
6. Demostrar que la ley asociativa de la suma y producto es válida en $D[x]$.
7. La «derivada formal» de $p(x)=a_0+a_1x+\ldots+a_nx^n$ se define como $p'(x)=a_1+2a_2x+\ldots+na_{n-1}x^{n-1}$. Demostrar que en cualquier dominio de integridad,

   a) $(cp)'=cp'$, b) $(p+q)'=p'+q'$,

   c) $(pq)'=pq'+p'q$, d) $(p^n)'=np^{n-1}p'$.

* 8. Si $p(y)$ y $q(x)$ son formas polinómicas con las indeterminadas y y x, mostrar que la substitución $y=q(x)$ da un polinomio $p(q(x))$. Para la derivada formal del Ejercicio 7, demostrar que $[p(q(x))]'=p'(q(x))\cdot q'(x)$.

* 9. Mostrar cómo, para un D dado, se construye un dominio de integridad $D(t)$ que consiste en todas las series infinitas «formales» de potencias con una indeterminada t, $a_0+a_1t+a_2t^2+\ldots$, con coeficientes a_i en D.

2. Funciones polinómicas

Como antes, sea D un dominio de integridad, y sea

$$f(x)=a_0+a_1x+\dots+a_mx^m$$

una forma polinómica en x sobre D. Si la indeterminada x se reemplaza por un elemento $c \epsilon D$, $f(x)$ deja de ser una simple expresión formal, y puede ser evaluada como un elemento definido $(a_0+a_1c+\dots+a_mc^m) \epsilon D$. En otras palabras, si se mira x como una variable independiente en el sentido del álgebra elemental, en lugar de considerarla un símbolo abstracto extraño a D, $f(x)$ se convierte en una *función:* «Si x es dado (como c), $f(x)$ resulta determinado (como $f(c)$)». Abstractamente, definiremos una «función» f de una variable, sobre D, como una regla que hace corresponder a cada elemento x en D, un «valor» $f(x)$, también en D. Diremos que dos funciones son *iguales* (simbólicamente, $f=g$), cuando sea $f(x)=g(x)$ para todo x. La *suma* $h=f+g$, la *diferencia* $h'=f-g$ y el *producto* $h^*=fg$ de dos funciones, se definen por la regla de que, para cualquier valor de x, $h(x)=f(x)+g(x)$, $h'(x)=f(x)-g(x)$ y $h^*=f(x)g(x)$. Función *constante* es aquella cuyo valor b es independiente de x. La función *idéntica* es la función f_0 tal, que $f_0(x)=x$ para todo x.

DEFINICIÓN. *Una función polinómica es una función que puede escribirse en la forma (1).*

Como las únicas reglas utilizadas para deducir las fórmulas (2) y (3) son válidas en cualquier dominio de integridad, serán válidas para cualquier valor c, en D, que se asigne a la indeterminada x (*). Esto significa que son identidades; por lo tanto, las sumas y productos de funciones polinómicas, se calculan por las fórmulas (2) y (3).

Por definición, cada forma (1) determina sólo una función polinómica, y cada función polinómica es determinada al menos por una de tales formas. Existe ciertamente una correspondencia pluriunívoca, que conserva sumas y productos, entre las formas poli-

(*) Precisamente éste es el secreto de la resolución de ecuaciones diciendo «sea x la cantidad desconocida»: las operaciones realizadas sobre x deben ser válidas para cualquier valor de x.

nómicas y las funciones polinómicas en cualquier dominio de integridad D que se considere.

Si estuviésemos seguros de que la correspondencia es biunívoca, sabríamos que se trata de un isomorfismo. Entonces, desde el punto de vista abstracto, podríamos olvidar la distinción entre formas polinómicas y funciones polinómicas. Desgraciadamente, no es éste el caso.

Para verlo, consideremos las dos formas distintas $f(x)=x^4+x$ y $g(x)=x^3+x^2$ sobre el campo J_3 de enteros módulo 3. Es fácil calcular los valores $f(0)=0$, $f(1)=2$, $f(2)=0$, y $g(0)=0$, $g(1)=2$, $g(2)=0$ para las funciones correspondientes (fig. 2). Aquí se ve que, por lo menos en este caso, formas diferentes determinan la misma función: la *igualdad* tiene, en efecto, un significado distinto para las funciones que para las formas.

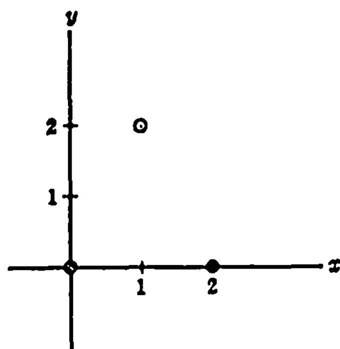


Figura 2

Vamos ahora a ver que no es accidental la suposición de que el dominio de los coeficientes sea finito, como en el ejemplo anterior. Demostraremos que no se puede construir un ejemplo análogo sobre el campo de los racionales. Pero antes de hacerlo, recordaremos algunas nociones elementales. El grado n de una forma no nula (1) es su exponente más elevado. Al término $a_n x^n$ de mayor grado se le llama término *principal*; su coeficiente a_n es el *principal* de (1); si $a_n=1$, el polinomio se llamará *mónico*.

TEOREMA 3. *Una forma polinómica* r(x) *de grado* n *sobre un dominio de integridad* D, *tiene a lo sumo* n *ceros en* D.

(Un cero de $r(x)$ significa una raíz de la ecuación $r(x)=0$, o sea, un elemento $a \in D$ tal, que $r(a)=0$.)

Demostración. Sea $r(x)=c_0+c_1x+c_2x^2+\ldots+c_nx^n$ $(c_n \neq 0)$. Para cualquier a tendremos, según el álgebra ordinaria,

$$\sum_{k=0}^{n} c_k x^k - \sum_{k=0}^{n} c_k a^k = \sum_{k=0}^{n} c_k (x^k - a^k) = \sum_{k=0}^{n} c_k[(x-a)(x^{k-1}+x^{k-2}a+\ldots+a^{k-1})]$$

Por lo tanto, $r(x)-r(a)=(x-a)s(x)$, en donde $s(x)$ es una forma polinómica de grado $n-1$. De aquí que, si a_1 es un cero de $r(x)$, será formalmente $r(x)=(x-a_1)s(x)$. Pero ahora se ve, por el Teorema 1, Capítulo I, que los ceros de $r(x)$ son a_1 y los de $s(x)$. Por inducción, $s(x)$ tiene a lo sumo $(n-1)$ ceros; por lo tanto, $r(x)$ tiene a lo más n ceros, c. q. d.

TEOREMA 4. *Si D es un dominio de integridad infinito, dos formas polinómicas sobre D son iguales si definen la misma función.*

Demostración. Como en (1), sean $p(x)$ y $q(x)$ dos formas dadas en la indeterminada x. Ambas determinan la misma función si $p(a)=q(a)$ para cualquier elemento a elegido en D; la conclusión que anunciamos es que $p(x)$ y $q(x)$ tendrán el mismo grado e iguales coeficientes homólogos. Utilizando la diferencia $r(x)=$ $=p(x)-q(x)$, esto quiere decir que $r(a)=c_0+c_1a+\ldots+c_na^n=0$ para todo $a \in D$, implica que $c_0=c_1=\ldots=c_n=0$. Esta conclusión se deduce del Teorema 3, pues, a menos que todos sus coeficientes fuesen nulos, el polinomio $r(x)$ podría anularse para n valores de x a lo sumo; mientras que, como D es infinito, son infinitos los valores a de x para los que $r(x)\neq 0$.

Así pues, si D es infinito, los conceptos de forma polinómica y función polinómica son equivalentes (dicho técnicamente: el anillo de funciones polinómicas es isomorfo con el de formas polinómicas). Otra manera de enunciar el Teorema 4 es decir que las formas (1) son un conjunto de formas canónicas para las funciones polinómicas, en el sentido de que cada función polinómica puede reducirse a una, y sólo a una, de las formas (1). Por otra parte, el Teorema 4 no se verifica si D es un dominio de integridad finito, con elementos a_1, a_2, ..., a_n. Por ejemplo, la forma polinómica mónica $(x-a_1)(x-a_2)\ldots(x-a_n)$, de grado n, determina en este caso la misma función que la forma 0.

Como cada sistema isomorfo con un dominio de integridad es asimismo un dominio de integridad, el Teorema 4 implica el siguiente

COROLARIO. *Las funciones polinómicas sobre un dominio de integridad infinito forman un dominio de integridad.*

Si D es un campo infinito, las formas racionales distintas definen funciones racionales diferentes, y las *funciones* racionales sobre D forman un campo.

(Nota. Una función racional no está definida en todos los puntos, sino únicamente en aquellos para los que el denominador es distinto de cero. Así pues, resulta definida en todos los puntos del campo D, excepto, quizá, en un número finito de ellos.)

EJERCICIOS

1. En el dominio de los enteros mód. 5, hallar una segunda forma polinómica que determine la misma función que $x^2 - x + 1$.
2. Mostrar que $x^2 \equiv 1$ (mód. 15) tiene cuatro ceros. ¿Por qué no contradice esto al Teorema 3?
3. a) Si $a_0, a_1, \ldots, a_n$, son $n+1$ elementos distintos de un campo F, demostrar que no puede haber dos formas polinómicas distintas $r(x)$ y $s(x)$ de grado $\leq n$ sobre D con $r(a_i) = s(a_i)$, $i = 0, 1, \ldots, n$.
   b) En a), discutir el grado de la forma polinómica $r(x) = (x - a_1)(x - a_2)\ldots(x - a_n)$, y calcular $r(a_0), r(a_1), \ldots, r(a_n)$.
   c) Mediante b), hallar un polinomio real, $s(x)$, de grado n, tal que $s(a_0) = 1$, $s(a_1) = \ldots = s(a_n) = 0$.
   d) Sean $c_0, \ldots, c_n$ elementos cualesquiera dados en F. Utilizando c), construir un polinomio $f(x)$ de grado $\leq n$, tal que $f(a_0) = c_0$, $f(a_1) = c_1$, ..., $f(a_n) = c_n$. (Ésta es la fórmula de interpolación de Lagrange.)
4. Hallar un polinomio cúbico $f(x) = a + bx + cx^2 + dx^3$ satisfaciendo a $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, $f(3) = 1$, considerando para ello a a, b, c, d como las incógnitas en cuatro ecuaciones, de las cuales la última es $a + 3b + 9c + 27d = 1$. (Este es el método de coeficientes indeterminado; compararlo con el método del Ejercicio 3.)
5. Utilizar la fórmula de interpolación del Ejercicio 3 para mostrar que cualquier función en cualquier dominio de integridad finito (tal como J_p) es una función polinómica.
6. * Sea D un dominio de integridad finito con n elementos $a_1, a_2, \ldots, a_n$. Sea $m(x)$ la forma polinómica $(x - a_1)(x - a_2)\ldots(x - a_n)$.
   a) Demostrar que si dos formas polinómicas $f(x)$ y $g(x)$ determinan la misma función, entonces $m(x)$ es divisor de la forma $f(x) - g(x)$. ¿Es cierta la recíproca?
   b) Calcular $m(x)$ para los dominios J_3 y J_5.
   c) Demostrar con el teorema del binomio para el desarrollo de $(1 + \ldots + 1)^p$, que $m(x) = x^p - x$ en el caso $D = J_p$.
7. Demostrar que sobre un campo infinito, formas racionales distintas determinan funciones distintas.
8. a) Si D y D′ son dominios isomorfos, demostrar que D[x] es isomorfo a D′[y], donde D[x] y D′[y] son los dominios de las formas polinómicas con las indeterminadas x e y sobre D y D, respectivamente.
   b) ¿Cómo son entre sí D(x) y D′(y)?
9. Si F es el campo de cocientes de un dominio D (Cap. II), demostrar que el campo D(x) es isomorfo con el campo F(x).

* 3. Divisores de cero y anillos conmutativos

Sea D un dominio de integridad finito y designemos por $D\{x\}$ el sistema de funciones polinómicas sobre D. Para todo $x \in D$, $f(x)+g(x)=g(x)+f(x)$, $0+f(x)=f(x)$, $1 \cdot f(x)=f(x)$, etc. Por tanto, la adición y la multiplicación son conmutativas, asociativas y distributivas, existen elementos idénticos para la adición y la multiplicación e inversos para la adición. En resumen, $D\{x\}$ satisface a todos los postulados requeridos para los dominios de integridad, excepto la ley de simplificación para la multiplicación. Ésta no se cumple, como hemos visto, porque existe un producto nulo tal como $(x-a_1)(x-a_2)...(x-a_n)$ cuyos factores son distintos de cero.

Análogamente, los enteros módulo m, siendo m compuesto, que vimos en el Capítulo I, § 10, satisfacen todos los postulados para un dominio de integridad excepto la ley relativa a los divisores de cero. Estos ejemplos, y otros análogos, nos sugieren el concepto de *anillo conmutativo*.

DEFINICIÓN. *Un anillo conmutativo es un conjunto con dos operaciones binarias, llamadas adición y multiplicación, tales que, además de ser conmutativas y asociativas,*

1) *la multiplicación es distributiva respecto de la adición;*

2) *existen un elemento idéntico para la adición y un opuesto para cada elemento.*

Un anillo conmutativo *con unidad* es un anillo en el que, además, existe un elemento, 1, idéntico para la multiplicación, de modo que $1 \cdot x = x$, para todo x. Así, $D\{x\}$ y J_{12} son anillos conmutativos con unidad (*). Recordemos que las reglas 1-9 del § 2, Capítulo I, se demostraron sin utilizar la ley de simplificación de la multiplicación; por lo tanto, son válidas en todo anillo conmutativo con unidad.

Otro ejemplo de anillo conmutativo con unidad nos lo proporciona el sistema D^* de *todas* las funciones sobre un dominio de integridad D, en el que la adición y la multiplicación hayan sido definidas como en el § 2. Existen divisores de cero en el dominio D^* de todas las funciones, aun siendo D un dominio de integridad de

(*) Un anillo más general (no conmutativo y sin elemento unidad) se verá en Capítulo XIII.

infinitos elementos. Así, si D es un dominio *ordenado* y si definimos $f(x)=|x|+x$ y $g(x)=|x|-x$, entonces $fg=h$ es $h=|x|^2-x^2=0$ para todo x, aunque $f\neq 0$, $g\neq 0$. Por lo demás, D^* tiene las restantes propiedades características de un dominio de integridad (Capítulo I, § 2). Se puede probar cada ley en D^* como la correspondiente ley en D, con el simple artificio de escribir «para todo x» en el segundo miembro. Así, $f(x)+g(x)=g(x)+f(x)$ para todo x, implica $f+g=g+f$. Si ahora definimos e como la función constante $e(x)=1$ para todo x, entonces $e(x)f(x)=1\cdot f(x)=f(x)$ para todo x y f, implica $ef=f$ para todo f. (Véase por qué la ley de simplificación de la multiplicación no puede probarse de este modo.) Como la ley de simplificación para la multiplicación no se ha utilizado en lo anterior, podemos enunciar :

TEOREMA 5. *Las funciones sobre un anillo conmutativo con unidad forman un anillo conmutativo con unidad.*

Definamos ahora (por analogía con subdominio) un *subanillo* de un anillo conmutativo A con unidad, como un subconjunto de A que contiene, con cada dos elementos f y g, también $f\pm g$ y fg.

Por el Teorema 1, el conjunto $D\{x\}$ de las funciones polinómicas, sobre un dominio de integridad D, es : 1.°, un subanillo del anillo D^* de todas las funciones sobre D, el cual, 2.°, contiene todas las funciones constantes y la función idéntica, y 3.°, está contenido en cualquier otro subanillo que cumple estas dos condiciones. En tal sentido, $D\{x\}$ es el subanillo de D^* engendrado por las funciones constantes y la función idéntica. Esto nos da una sencilla caracterización algebraica del concepto de función polinómica.

EJERCICIOS

1. a) Demostrar que sólo hay cuatro funciones diferentes en el campo J_2 de enteros mód. 2; escribir las tablas de adición y multiplicación para este anillo de funciones.
   b) Expresar cada una de estas funciones como una función polinómica.
   c) ¿Es este anillo de funciones isomorfo con el anillo de enteros mód. 4?
2. ¿Cuántas funciones diferentes hay sobre el anillo J_n de los enteros módulo n?
3. ¿Son los siguientes conjuntos de funciones anillos conmutativos con unidad?
   a) Todas las funciones f sobre un dominio D para las que $f(0)=0$;
   b) Todas las funciones f sobre D con $f(0)=f(1)$;

c) Todas las funciones f sobre D con $f(0) \neq 0$;
d) Todas las funciones f sobre R* (números reales) con $-7 \leq f(x) \leq 7$ para todo x;
e) Todas las funciones f acotadas sobre R* (f es *acotada* si existe un número real M tal que $-M \leq f(x) \leq M$ para todo x);
f) Todas las f sobre R*. con $f(x+1)=f(x)$ para todo x (funciones periódicas).

4. Construir dos anillos conmutativos de funciones no incluídas en los ejemplos del Ejercicio 3.
5. Sea D* definido como en el texto. Demostrar la ley asociativa para las sumas y productos en D*.
6. a) Si D y D' son dominios isomorfos, demostrar que $D\{x\}$ y $D'\{x\}$ son isomorfos.
b) ¿Cómo son entre sí D* y (D')*?

*7. Discutir, desde el punto de vista del Análisis, el subanillo del anillo de las funciones reales engendrado por las funciones constantes y sen x; y por las constantes, sen x y cos x.

4. Polinomios de varias variables. Automorfismos

La discusión que en los anteriores párrafos hemos aplicado a funciones (formas) polinómicas de una sola variable (indeterminada) x, puede extenderse en muchos aspectos, sin dificultad, al caso de varias variables (indeterminadas) $x_1, x_2, \ldots, x_n$.

Definición. *Una función polinómica de variables* $x_1, x_2, \ldots, x_n$, *sobre un dominio de integridad* D, *es una función que puede construirse por adición, sustracción y multiplicación, a partir de las funciones constantes* $f(x_1, x_2, \ldots, x_n)=c$ *y de las* n *funciones idénticas* $f(x_1, x_2, \ldots, x_n)=x_i$ $(i=1, 2, \ldots, n)$. *Una forma polinómica sobre* D, *en las indeterminadas* $x_1, x_2, \ldots, x_n$, *puede definirse por recurrencia como una forma en* x_n *sobre el dominio* $D[x_1, x_2, \ldots, x_{n-1}]$ *de las formas polinómicas en* $x_1, x_2, \ldots, x_{n-1}$ (*brevemente:* $D[x_1, x_2, \ldots, x_n]=D[x_1, x_2, \ldots, x_{n-1}][x_n]$).

Así, en el caso de dos variables x, y, una de tales formas será $p(x, y)=(3+x^2)+0\cdot y+(2x-x^3)y^2$ —con frecuencia se escribe en la forma más cómoda $3+x^2+2xy^2-x^3y^2$.

Un corolario del Teorema 4, por inducción sobre n, es

Teorema 6. *Una función polinómica en* $x_1, x_2, \ldots, x_n$, *puede expresarse de un solo modo como forma polinómica, si* D *es infinito. Sea* D *infinito o no lo sea,* $D[x_1, x_2, \ldots, x_n]$ *es siempre un dominio de integridad.*

DEFINICIÓN. *Un isomorfismo de un dominio de integridad* D *consigo mismo, se llama un automorfismo de* D.

De esta definición se desprende que cada permutación de los subíndices dará lugar a un automorfismo del anillo conmutativo $D\{x_1, x_2, \ldots, x_n\}$ de las funciones polinómicas de n variable. De aquí, por el Teorema 6, que si D es infinito, análogos automorfismos valen para las formas polinómicas (cuya definición no es simétrica en las variables). Vamos a ver ahora que este resultado es cierto sin esta restricción sobre D.

TEOREMA 7. *Cada permutación de los subíndices determina un automorfismo distinto en* $D[x_1, x_2, \ldots, x_n]$.

Consideremos el caso de dos indeterminadas x, y. Cada forma $p(y, x) = \sum_i (\sum_j a_{ij} y^j) x^i$ de $D[y, x]$ puede ser reordenada por las leyes distributiva, conmutativa y asociativa en $D[y, x]$, para dar una expresión de la forma $p(y, x) = \sum_j (\sum_i a_{ij} x^i) y^j$. Este resultado puede interpretarse como si fuese un polinomio $p'(x, y)$ en el dominio $D[x, y]$ (x primero, después y). La correspondencia $p(y, x) \leftrightarrow$ $\leftrightarrow p'(x, y)$ así establecida es biunívoca; cada conjunto de coeficientes a_{ij} no nulos, corresponde a un solo elemento de $D[y, x]$ y a uno solo de $D[x, y]$. Finalmente, como las reglas (2)-(3) para adición y multiplicación pueden deducirse de los postulados para un dominio de integridad, cosa que son tanto $D[y, x]$ como $D[x, y]$, resulta que la correspondencia conserva sumas y productos, c. q. d.

El caso de n indeterminadas puede tratarse análogamente, con una notación general más complicada, o deducirse por inducción partiendo del caso de dos variables.

Resulta así que en realidad $D[x_1, x_2, \ldots, x_n]$ depende *simétricamente* de $x_1, x_2, \ldots, x_n$. Esto nos incita a buscar una definición de $D[x_1, x_2, \ldots, x_n]$ en la cual esta simetría sea patente. Indicaremos la marcha a seguir en el caso $n = 2$, para el dominio $D' = D[x, y]$. En primer lugar, D' es engendrado por x, y, y los elementos de D (todo elemento de D' puede obtenerse de x, y, y D por repetidas sumas y productos); en segundo lugar, los elementos x e y son *indeterminadas simultáneas* sobre D (o bien, *algebraicamente independientes* sobre D). Con esto queremos expresar que una suma finita $\sum_{i,j} a_{ij} x^i y^j$ con coeficientes a_{ij} en D, será

cero si, y sólo si, todos los coeficientes a_{ij} son nulos. Estas dos propiedades bastan para determinar el dominio $D[x, y]$ de una manera simétrica (ver el siguiente Ejercicio 9).

EJERCICIOS

1. Representar como polinomios en y con coeficientes en $D[x]$:
   a) $p(x, y) = y^3x + (x^2 - xy)^2$; b) $q(x, y) = (x+y)^3 - 3yx(x^2+x-1)$.
2. Contar el número de funciones posibles de dos variables x, y, sobre el dominio J_2.
3. Reordenar las siguientes expresiones como polinomios en x, cuyos coeficientes sean polinomios en y (como en la demostración del Teorema 7): $(3x^2+2x+1)y^3 + (x^4+2)y^2 + (2x-3)y + x^4 - 3x^2 + 2x$.
4. Sea D un dominio de integridad. Demostrar que la correspondencia que transporta cada $p(x)$ sobre $p(-x)$ es un automorfismo de $D[x]$. ¿Lo es también para $D\{x\}$?
5. ¿Es la correspondencia $p(x) \to p(x+c)$, donde c es constante, un automorfismo de $D[x]$? Ilustrarlo con ejemplos numéricos.
6. Si F es un campo, demostrar que la correspondencia $p(x) \to p(ax)$ es un automorfismo de $F[x]$ para cualquier constante $a \neq 0$.
7. ¿Existen otros automorfismos de $D[x, y]$ aparte los descritos en el Teorema 7?
8. Demostrar el Teorema 6. a) para $n=2$; b) para cualquier n.
9. a) Demostrar con detalle que el dominio $D[x, y]$ (primero x, después y) es igualmente engendrado sobre D por dos «indeterminadas simultáneas», x e y.
   b) Sean D′ y D″ dos dominios engendrado cada uno sobre D por dos indeterminadas simultáneas x', y' y x'', y'', respectivamente. Demostrar que D′ es isomorfo con D″ bajo una correspondencia que representa x' sobre x'', y' sobre y'', y cada elemento de D sobre sí mismo.
   c) Utilizando a) y b), dar otra demostración del Teorema 7.

5. Divisibilidad

La célebre descomposición factorial $x^2 - 1 = (x-1)(x+1)$ muestra que $x - 1$ es un *factor* o *divisor* del polinomio $x^2 - 1$. En cambio, el polinomio $x^2 + 1$ no puede descomponerse en factores polinómicos con coeficientes racionales, excepto de maneras triviales, tales como $x^2 + 1 = (2x^2+2)\cdot 1/2$. Los polinomios como éste, sin divisores propios sobre un campo R, se llaman *irreducibles* en tal campo. Los términos «divisor» e «irreducible» son análogos, en el dominio de los polinomios $R[x]$, a los términos «divisor» y «primo» en el dominio de los enteros. Nuestro principal objeto, en los §§ 5-7, es demostrar los teoremas análogos al teorema fun-

damental de la Aritmética: probar que todo polinomio puede ser descompuesto de un solo modo en factores polinómicos «irreducibles» (salvo descomposiciones triviales).

Para revelar dicha analogía, y también para evitar repeticiones, comenzaremos por definir algunos conceptos válidos en un dominio de integridad D en general, ya sea el dominio de los polinomios $R[x]$, el dominio J de los enteros u otro dominio cualquiera.

Un elemento a de D es divisible por b (simbólicamente, $b \mid a$) si existe en D algún elemento c tal, que $a = cb$. Dos elementos a y b son *asociados* si se verifican las dos condiciones $b \mid a$ y $a \mid b$. Todo elemento asociado del 1 será llamado *unidad* en el campo. Como $1 \mid a$ para todo a, un elemento u será unidad en D si existe en D su inverso respecto a la multiplicación, u^{-1}, con $1 = uu^{-1}$. Si a y b son asociados, $a = cb$ y $b = c'a$, luego $a = cc'a$. Por la ley de simplificación queda $1 = cc'$, y por tanto, c y c' son unidades. Recíprocamente, $a = ub$ es asociado de b si u es una unidad. De esto deducimos que, para que dos elementos sean asociados, es necesario y suficiente que cada uno pueda obtenerse multiplicando el otro por un factor unidad.

EJEMPLO 1. En un campo, todo $a \neq 0$ es unidad.

EJEMPLO 2. En el dominio J de los enteros, las unidades son ± 1; en consecuencia, los asociados de cualquier entero a son $\pm a$.

EJEMPLO 3. En un dominio $D[x]$ de polinomios en una indeterminada x, el grado de un producto $f(x) \cdot g(x)$ es la suma de los grados de los factores. Por lo tanto, si un elemento $b(x)$ tiene polinomio inverso $a(x)$, como $a(x) \cdot b(x) = 1$, tal polinomio $b(x)$ deberá ser de grado cero. Pero un polinomio constante b tiene un inverso si b tiene ya un inverso en D, y sólo en este caso. Así pues, las unidades de $D[x]$ son las mismas unidades de D.

Si F es un campo, las unidades del dominio de polinomios $F[x]$ son exactamente las constantes no nulas de F; así que dos polinomios $f(x)$ y $g(x)$ son asociados en F cuando cada uno de ellos es el producto del otro por una constante.

EJEMPLO 4. En el dominio $J[\sqrt{2}]$ de todos los números $a + b\sqrt{2}$ (a, b, enteros), $(a + b\sqrt{2})(x + y\sqrt{2}) = 1$ implica $x = a/(a^2 - 2b^2)$, $y = -b/(a^2 - 2b^2)$; éstos son enteros cuando $a^2 - 2b^2 = \pm 1$. Así resulta que $1 \pm \sqrt{2}$ y $3 \pm 2\sqrt{2}$ son unidades en $J[\sqrt{2}]$, mientras que $2 + \sqrt{2}$ no lo es.

Un elemento b de un dominio de integridad arbitrario D, es divisible por todos sus asociados y por todas las unidades. Éstos se llaman divisores «impropios» de b. Un elemento que no sea unidad y que no tenga divisores propios, se llama *primo* o *irreducible* en D.

Ejemplo 5. Sobre cualquier campo F, un polinomio lineal $ax+b$, con $a \neq 0$, es irreducible, porque sus factores o son constantes (unidades), o producto de él mismo por constantes (asociados).

Ejemplo 6. El polinomio x^2+4 es irreducible en el campo de los racionales. Para verlo, supongamos que fuese $x^2+4=(x+a)(x+b)$. Poniendo $x=-b$, esto daría $(-b)^2+4=(-b+a)(-b+b)=0$, luego $(-b)^2=-4$. Esto es evidentemente imposible, pues un cuadrado no puede ser negativo. Como este razonamiento es válido para cualquier campo ordenado, deducimos que x^2+4 es también irreducible en el campo real o en cualquier otro campo ordenado. Esto no quiere decir que x^2+4 sea irreducible en todos los campos. Así, sobre el de los números complejos (Cap. V) se tiene la reducción $x^2+4=(x-2i)(x+2i)$. Destaquemos, pues, el hecho de que la divisibilidad no pasa como herencia de dominio a subdominio. Así, 2 no es unidad en el dominio de los enteros, paro en el dominio de los números racionales tiene por inverso $\frac{1}{2}$ y, así, es unidad.

EJERCICIOS

1. En cualquier dominio de integridad D, demostrar:
   a) La relación «$b \mid a$» es reflexiva y transitiva;
   b) Si $c \neq 0$, es $b \mid a$ si, y sólo si, $bc \mid ac$;
   c) Dos elementos cualesquiera tienen un divisor común y un múltiplo común;
   d) Si $a \mid b$ y $a \mid c$, también será $a \mid (b \pm c)$.
2. Demostrar las siguientes propiedades de las unidades en cualquier dominio:
   a) El producto de dos unidades es una unidad;
   b) Una unidad u de D divide a cualquier elemento de D;
   c) Si c divide a cualquier x en D, c es una unidad.
3. En cualquier dominio de integridad, indiquemos que «a es asociado de b» utilizando el signo «$a \sim b$». Demostrar que
   a) Si $a \sim b$, será $c \mid a$ si, y sólo si, $c \mid b$;
   b) Si $a \sim b$, será $a \mid c$ si, y sólo si, $b \mid c$;
   c) Si $a \mid c$ cuando, y sólo cuando, $b \mid c$, será $a \sim b$;
   d) Si p es primo y $p \sim q$, también q será primo.
4. Demostrar que si $a \sim a'$ y $b \sim b'$, también $ab \sim a'b'$. (En cambio, en general, no es $a+b \sim a'+b'$.)

5. Demostrar la «ley de simplificación generalizada»: si $ax \sim by$, $a \sim b$, y $a \neq 0$, entonces $x \sim y$.
6. ¿Es x^2+1 irreducible sobre J_3? ¿Y sobre J_5? Lo mismo para x^2+x+2.
7. Hallar un campo finito sobre el cual x^2-2 sea: a) reducible; b) irreducible.
8. ¿Cuáles de los siguientes polinomios son irreducibles sobre el campo $R^{\#}$ de los números reales?

$$x^2+9, \quad x^3+8, \quad x^4+1, \quad x^4+5x^2+1, \quad x^4-2x+1.$$

9. Enumerar todos los asociados de x^2+2x-1 en $J_3[x]$.
10. Hallar todas las unidades en el dominio $D[x, y]$ de los polinomios con dos indeterminadas.
11. ¿Para que elementos a de un dominio de integridad D es un automorfismo de $D[x]$ la correspondencia $p(x) \to p(ax)$?
12. Hallar todas las unidades en el dominio D que consiste en todos los números racionales m/n con m y n enteros, no siendo n divisible por 7.
13. Para $\alpha = a + b\sqrt{2}$, definiremos $N(\alpha) = a^2 - 2b^2$. Demostrar:
    a) $N(\alpha\alpha') = N(\alpha)N(\alpha')$, y
    b) Si α es una unidad en $J[\sqrt{2}]$, será $N(\alpha) = \pm 1$.
14. Sea $J[\sqrt{5}]$ el dominio de todos los números $\alpha = a + b\sqrt{5}$ (a, b enteros), y sea $N(\alpha) = a^2 - 5b^2$.
    a) Demostrar que $9+4\sqrt{5}$ es una unidad del dominio. (Cfr. Ejerc. 13.)
    b) Demostrar que $1-\sqrt{5}$ y $3+\sqrt{5}$ son asociados, pero no unidades.
    c) Mostrar, generalmente, que α es una unidad si, y sólo si, $N(\alpha) = \pm 1$.
    d) Si $N(\alpha)$ es primo en J, demostrar que α es primo en $J[\sqrt{5}]$.
    e) Demostrar que $4+\sqrt{5}$ y $4-\sqrt{5}$ son primos.
    f) Demostrar que 2 y $3+\sqrt{5}$ son primos. (*Sugerencia*: $x^2 \equiv 2$ (mód. 5) es imposible para $x \in J$.)
    g) Utilizar la igualdad $2 \cdot 2 = (3+\sqrt{5}) \cdot (3-\sqrt{5})$ para mostrar que $J[\sqrt{5}]$ no es un dominio con factorización única (§ 8).

6. Algoritmo de la división

En Álgebra elemental se aprendió a dividir un polinomio $b(x)$ por otro $a(x)$, obteniendo un cociente $q(x)$ y un resto $r(x)$ de grado menor que el divisor $a(x)$. Vamos a ver ahora que este algoritmo, que se estableció con coeficientes racionales, es válido para polinomios con coeficientes de cualquier campo.

TEOREMA 8. *Si* F *es un campo, y* $a(x) \neq 0$ *y* $b(x)$ *son dos polinomios cualesquiera sobre* F, *podremos hallar dos polinomios* $q(x)$ *y* $r(x)$ *sobre* F *tales, que*

$$b(x) = q(x)a(x) + r(x) \tag{4}$$

siendo $r(x)$ *o bien nulo o bien de grado menor que* $a(x)$.

Demostración. Basta eliminar sucesivamente los términos de mayor grado del dividendo $b(x)$, restando de él los productos del divisor $a(x)$ por monomios cx^k elegidos convenientemente. Si $a(x) = a_0 + a_1x + \ldots + a_mx^m$ $(a_m \neq 0)$ y $b(x) = b_0 + b_1x + \ldots + b_nx^n$ $(b_n \neq 0)$, y si el grado n de $b(x)$ no es menor que el grado m de $a(x)$, podremos formar la diferencia

$$(5) \quad b_1(x) = b(x) - (b_n/a_m)x^{n-m}a(x) = 0 \cdot x^n + (b_{n-1} - a_{m-1}b_n/a_m)x^{n-1} + \ldots,$$

la cual será de grado menor que n o cero. Podemos repetir este proceso hasta que el grado del resto sea menor que m.

Hablando propiamente, este razonamiento emplea el segundo principio de inducción, formulado en § 5, Capítulo I. Sea m el grado de $a(x)$. Cualquier polinomio de grado $n < m$ tiene una representación $b(x) = 0 \cdot a(x) + b(x)$, con un cociente $q(x) = 0$. Para un polinomio $b(x)$ de grado $n \geqslant m$, trasponiendo (5) se tiene

$$(6) \qquad b(x) = b_1(x) + (b_n/a_m)x^{n-m}a(x),$$

donde el grado k de $b_1(x)$ es menor que n, a menos que $b_1(x) = 0$. Por el segundo principio de inducción, podemos admitir la descomposición (4) para todo $b(x)$ de grado $k < n$, así que tenemos

$$(7) \qquad b_1(x) = q_1(x)a(x) + r(x),$$

siendo el grado de $r(x)$ menor que m, salvo que sea $r(x) = 0$.

Sustituyendo (7) en (6) obtenemos la igualdad (4) que deseábamos :

$$b(x) = [q_1(x) + (b_n/a_m)x^{n-m}]a(x) + r(x).$$

En particular, si el polinomio $a(x) = x - c$ es mónico y lineal, entonces el resto $r(x)$ en (4) es una constante, $r(x) = b(x) - (x - c)q(x)$. Haciendo $x = c$ esta igualdad dará $r = b(c) - 0q(c) = b(c)$. De aquí :

Corolario 1. *El resto de dividir un polinomio* p(x) *por el binomio* x — c, *es* p(c). *(Teorema del resto.)*

En particular, $p(c)$ será cero cuando $x - c$ sea un factor de $p(x)$. Luego :

Corolario 2. *Cada raiz* c *de* p(x) *proporciona un factor lineal,* x — c, *de* p(x) ; *y reciprocamente.*

EJERCICIOS

1. Demostrar que $q(x)$ y $r(x)$ están determinados unívocamente por $a(x)$ y $b(x)$ en (4).
2. Calcular $q(x)$, $r(x)$, si $b(x)=x^4-x^3+3x-5$ y $a(x)=x^2+7$.
3. El mismo Ejercicio 2, si $a(x)$ es, respectivamente, $x-2$, $x+2$, x^2+x-1.
4. a) Resolver el Ejercicio 2 en el campo J_5.
   b) Resolver el Ejercicio 3 en el campo J_3.
5. ¿Es x^3+x^2+x+1 divisible por x^2+3x+2 en cualquiera de los dominios J_3, J_5, J_7?
6. Hallar todos los anillos posibles J_n en los que $x^3-10x+12$ sea divisible por x^2+2.
7. Hallar todos los polinomios cuadráticos irreducibles sobre el campo de los enteros mód. 5.
8. Hallar todos los polinomios cúbicos irreducibles sobre el campo de los enteros mód. 3.
9. a) Si un polinomio $f(x)$ sobre cualquier dominio tiene $f(a)=0=f(b)$, donde $a \neq b$, demostrar que $f(x)$ es divisible por $(x-a)(x-b)$.
   b) Generalizar este resultado.
10. Encontrar el teorema que sustituye al 8 si el campo F se reemplaza por un dominio de integridad D.
11. a) En la aplicación del segundo principio de inducción al algoritmo de división, ¿qué significado tiene $P(n)$? (Ver Cap. I, § 5.)
    b) Demostrar el algoritmo de división, utilizando el principio de buena ordenación en vez del segundo principio de inducción.
12. Demostrar que si $a_0+a_1x+a_2x^2+\ldots+a_nx^n$ es irreducible, también lo es $a_n+a_{n-1}x+a_{n-2}x^2+\ldots+a_0x^n$.

*13. a) Definir el m. c. d. y el m. c. m. en un dominio de integridad arbitrario D.
   b) Demostrar que dos m. c. d. de los mismos números a y b de D, son asociados.

7. Teorema de unicidad de la descomposición factorial

En este § 7 consideraremos la descomposición factorial en el dominio $F[x]$, de formas polinómicas en una indeterminada x sobre un *campo* F. El resultado principal es que la descomposición en factores irreducibles (primos) es única, siendo la demostración, esencialmente, una repetición de la del análogo teorema fundamental de la aritmética (Cap. I).

TEOREMA 9. *Todo subconjunto no vacío de polinomios de* F[x] *que contenga a la suma y diferencia de dos cualesquiera de ellos, así como a todos los múltiplos de uno cualquiera, consiste, o bien (1) en el cero sólo, o bien (2) en el conjunto de los múltiplos*

q(x)a(x) *de uno de sus elementos* a(x), *distinto de cero y con grado mínimo.*

Un conjunto de polinomios que tenga las dos propiedades enunciadas en la hipótesis se llama un *ideal* C del dominio de polinomios $F[x]$; así que el teorema afirma que cada ideal de $F[x]$ está constituído por los múltiplos de un elemento particular $a(x)$ o bien es cero. Los ideales en dominios generales los estudiaremos en el Capítulo XIII.

Demostración. Si C no es cero, contendrá un polinomio $a(x)$ no nulo de grado mínimo $d(a)$, y con $a(x)$ estarán todos sus múltiplos $q(x)\cdot a(x)$. En tal caso, si $b(x)$ es *cualquier* polinomio de C, por el Teorema 8, algún $r(x)=b(x)-q(x)a(x)$ tiene grado menor que $d(a)$. Pero, por hipótesis, C contiene a $r(x)$, y por construcción, no hay en C ningún polinomio *distinto de cero* con grado menor que $d(a)$. De aquí $r(x)=0$ y $b(x)=q(x)a(x)$, lo que prueba el teorema.

Sean ahora $a(x)$ y $b(x)$ dos polinomios, y consideremos el conjunto C de todas las combinaciones lineales de ambos, $s(x)a(x)+$ $+t(x)b(x)$, con coeficientes polinomios $s(x)$ y $t(x)$. Este conjunto C no es vacío, y contiene las sumas, diferencias y múltiplos de sus elementos, ya que (en notación abreviada)

$$(sa+tb)\pm(s'a+t'b)=(s\pm s')a+(t\pm t')b,$$

$$q(sa+tb)=(qs)a+(qt)b.$$

Luego, por el Teorema 9, el ideal C de estas combinaciones lineales estará formado por los múltiplos de cierto polinomio $d(x)$ de grado mínimo en C.

Este $d(x)$ dividirá a $a(x)=1\cdot a(x)+0\cdot b(x)$, y a $b(x)=0\cdot a(x)+$ $+1\cdot b(x)$, y será divisible por cualquier divisor común de $a(x)$ y $b(x)$, ya que $d(x)=s(x)a(x)+t(x)b(x)$. Nuestra conclusión es la siguiente:

TEOREMA 10. *En* F[x], *dos polinomios cualesquiera* a *y* b *tienen un «máximo común divisor»* d *tal, que (1.°)* d|a *y* d|b, *(2.°)* c|a *y* c|b *implica* c|d. *Además, (3.°)* d *es una combinación lineal,* d = sa + tb, *de* a *y* b.

Observamos que el algoritmo de Euclides, explicado detalladamente en el Capítulo I, § 7, puede utilizarse para el cálculo de d

a partir de a y b. (La razón es que el algoritmo de la división nos permite calcular explícitamente los restos.)

Además, si d satisface (1.°), (2.°) y (3.°), igual les sucederá a los asociados de d. Hagamos notar que (1.°) y (2.°) implican (3.°).

El m. c. d. $d(x)$ es único, salvo factores unidad, porque si d y d' son dos máximos comunes divisores de a y b, entonces por (1.°) y (2.°) $d \mid d'$ y $d' \mid d$, así que d y d' serán asociados. Recíprocamente, si d es un m. c. d., lo mismo le sucederá a cualquier asociado de d. A veces conviene considerar como m. c. d. al único polinomio mónico asociado de d.

Los polinomios $a(x)$ y $b(x)$ se llaman primos entre sí si sus máximos comunes divisores son una unidad y sus asociadas. Esto significa que la condición necesaria y suficiente para que dos polinomios sean primos entre sí, es que sus factores comunes sean las constantes de F (las unidades en el dominio $F[x]$).

TEOREMA 11. *Si* p(x) *es irreducible*, p(x) | a(x)b(x) *implica que o bien* p(x) | a(x), *o bien* p(x) | b(x).

Demostración. Puesto que $p(x)$ es irreducible, el m. c. d. de $p(x)$ y $a(x)$ será $p(x)$ o la unidad 1. En el primer caso, $p(x) \mid a(x)$; en el segundo, podemos escribir $1 = s(x)p(x) + t(x)a(x)$, y entonces,

$$b(x) = b(x) \cdot 1 = s(x)p(x)b(x) + t(x)a(x)b(x).$$

Como $p(x)$ divide el producto $a(x)b(x)$, dividirá a los dos sumandos del segundo miembro, luego dividirá al primer miembro $b(x)$, c. q. d.

TEOREMA 12. *Todo polinomio* a(x) *no constante en* F[x] *puede expresarse como el producto de una constante* c *por otros polinomios mónicos irreducibles. Esta descomposición es única, salvo el orden de los factores.*

Primeramente, tal descomposición es posible. Si $a(x)$ es una constante o es irreducible, el teorema es trivial. En otro caso, $a(x)$ es producto de factores de menor grado, $a(x) = b(x)b'(x)$. Por el segundo principio de inducción, podemos suponer que

$$b(x) = cp_1(x) \ldots p_m(x), \qquad b'(x) = c'p_1'(x) \ldots p_n'(x).$$

De aquí que $a(x) = cc'p_1(x) \ldots p_m(x)p_1'(x) \ldots p_n'(x)$, en donde cc' es una constante, y los $p_i(x)$ y $p_j'(x)$ son polinomios mónicos irreducibles.

Para probar la unicidad supongamos que $a(x)$ tuviese dos de estas descomposiciones:

$$a(x)=cp_1(x)\dots p_m(x)=c'q_1(x)\dots q_n(x).$$

Fácilmente vemos que $c=c'$, pues sería el coeficiente del término principal de $a(x)$ (ya que éste es el producto de los términos principales de los factores). Además, puesto que $p_1(x)$ divide a $c'q_1(x)\dots q_n(x)=a(x)$, debe, por el Teorema 11, dividir a algún factor no constante $q_i(x)$; como $q_i(x)$ es irreducible, el cociente $q_i(x)/p_1(x)$ debe ser constante; y como $p_1(x)$ y $q_i(x)$ son los dos mónicos, esta constante es 1. De aquí que $p_1(x)=q_i(x)$. Dividiendo por esta igualdad queda $p_2(x)\dots p_m(x)$ igual al producto de las $q_k(x)$ $(k\neq i)$, y tiene menor grado que $a(x)$. Ahora bien, aplicando de nuevo el segundo principio de inducción, las $p_j(x)$ $(j\neq 1)$ y $q_k(x)$ $(k\neq i)$ son iguales a pares, lo que completa la demostración.

Un corolario es (ver Cap. I, §8, último párrafo) que el exponente e_i con que aparece cada polinomio irreducible mónico $p_i(x)$ factor de $a(x)$, viene unívocamente determinado por $a(x)$ y es el mayor valor de e tal, que $p_i(x)^e \mid a(x)$.

Si un polinomio $a(x)$ se descompone en factores irreducibles $p_i(x)$ no necesariamente mónicos, las descomposiciones posibles no son absolutamente idénticas. Sin embargo, cada factor $p_i(x)$ dividido por el coeficiente de su término principal da un (único) factor mónico irreducible, y, por lo tanto, es asociado de este irreducible en $F[x]$. Vemos, pues, que dos de tales descomposiciones pueden referirse una a otra, sin más que reordenar los factores y reemplazarlos por asociados convenientes. Esto se resume diciendo que la descomposición de un polinomio en factores irreducibles es única, *salvo el orden y los factores unidad* (o salvo el orden y la sustitución de factores por otros asociados).

EJERCICIOS

1. a) Describir el algoritmo de Euclides para dos polinomios dados $a(x)$ y $b(x)$, y demostrar que el último resto no nulo del algoritmo es el m. c. d. de $a(x)$ y $b(x)$.
   b) ¿Puede utilizarse el algoritmo de Euclides para representar $d(x)$ en la forma $d=sa+tb$? En caso afirmativo, decir cómo.
2. a) Hallar el m. c. d. de x^3-1 y $x^4+x^3+2x^2+x+1$.
   b) Expresar este m. c. d. como una combinación $d(x)=s(x)a(x)+t(x)b(x)$ de los polinomios dados. (*Observación*: Los coeficientes pueden no ser enteros.)

3. Lo mismo para $2x^3+6x^2-x-3$, $x^4+4x^3+3x^2+x+1$.
4. Hacer el Ejercicio 3 suponiendo que los polinomios tienen coeficientes en J_3.
5. Demostrar que x^3+x+1 es irreducible módulo 5. (*Sugerencia*: Comprobar con todos los factores lineales posibles.)
6. Hallar la descomposición de los siguientes polinomios en J_3:

   a) x^2+x+1, b) x^3+x+2, c) $2x^3+2x^2+x+1$,

   d) x^4+x^3+x+1, *e) x^4+x^3+x+2.

7. Enumerar (prescindiendo de los asociados) todos los divisores de x^4-1 en el dominio de polinomios con coeficientes reales, probando que cualquier divisor de x^4-1 es asociado a uno de la lista presentada.
8. Lo mismo para x^6-1, x^8-1.
9. Descomponer el polinomio x^4-5x^2+6 en factores irreducibles, sobre el campo de los racionales, sobre el campo $R(\sqrt{2})$ de Capítulo II, § 1, y también sobre el campo real.
10. Demostrar que cualquier conjunto finito de polinomios sobre un campo tiene un m. c. d. que es una combinación lineal de los polinomios dados.
11. a) Demostrar que el conjunto de todos los múltiplos comunes de dos polinomios dados sobre un campo es un ideal.
    b) Deducir que dos polinomios tienen un m. c. m.; ilustrarlo hallando el mínimo común múltiplo de x^2+3x+2 y $(x+1)^2$.
12. Si un polinomio dado $p(x)$ sobre F, tiene la propiedad que $p(x) \mid a(x)b(x)$ implica siempre que, o bien $p(x) \mid a(x)$, o bien $p(x) \mid b(x)$, demostrar que $p(x)$ es irreducible sobre F.
13. Si $p(x)$ es un polinomio dado, tal que cualquier otro es primo con $p(x)$ o divisible por $p(x)$, demostrar que $p(x)$ es irreducible.
14. Si $m(x)$ es igual a una potencia de un polinomio irreducible, mostrar que $m(x) \mid a(x)b(x)$ implica que $m(x) \mid a(x)$ o $m(x) \mid [b(x)]^e$ para algún e.
15. Si $h(x)$ es primo con $f(x)$ y $g(x)$ simultáneamente, demostrar que $h(x)$ es primo con $f(x)g(x)$.
16. Si $h(x) \mid f(x)g(x)$ y $h(x)$ es primo con $f(x)$, será $h(x) \mid g(x)$.
17. Si $f(x)$ y $g(x)$ son primos entre sí en $F[x]$, y si F es un subcampo de K, demostrar que $f(x)$ y $g(x)$ son primos relativos también en $K[x]$.

*18. Si dos polinomios con coeficientes racionales tienen una raíz real común, demostrar que tienen un divisor común con coeficientes racionales el cual no es una constante.

19. Las siguientes condiciones dan ciertos conjuntos de polinomios con coeficientes racionales. ¿Cuáles de estos conjuntos son ideales? Cuando el conjunto sea un ideal, hallar en él un polinomio de grado mínimo.
    a) Todo $b(x)$ con $b(3)=b(5)=0$;
    b) Todo $b(x)$ con $b(3)\neq 0$ y $b(2)=0$;
    c) Todo $b(x)$ con $b(\sqrt{2})=0$;
    d) Todo $b(x)$ con $b(2)=b'(2)=0$, donde $b'(x)$ es la derivada formal de $b(x)$, definida en Ejercicio 7, § 1;
    e) Todo $b(x)$ con $b(3)=0$, $b(6)=b(7)$;
    f) Todo $b(x)$ tal, que alguna de sus potencias sea divisible por $(x+1)^4(x+2)$.

10. Sea S un conjunto de polinomios sobre F que contenga a la diferencia de dos cualesquiera de sus elementos, y que si contiene a $b(x)$ contenga también a $xb(x)$ y a $ab(x)$ para cualquier constante a en F. Demostrar que S es un ideal.

8. Otros dominios con descomposición factorial única

Consideremos el dominio $R[x, y]$ de formas polinómicas con dos indeterminadas sobre el campo racional R. Los únicos divisores comunes de $a(x, y)=x$ y $b(x, y)=y^2+x$ son 1 y sus asociados (unidades), pero no pueden existir dos polinomios $s(x, y)$, $t(x, y)$ tales, que $s(x, y)x+t(x, y)(y^2+x)=1$, pues nunca el primer miembro tendrá un término independiente no nulo, sean cualesquiera s y t. De modo semejante, en el dominio $J[x]$ de los polinomios con coeficientes enteros, es m. c. d. $(2, x)=1$, pues no es posible que $s(x)\cdot 2+t(x)\cdot x=1$. Por lo tanto, el Teorema 10 no es válido en estos dominios.

Sin embargo, se puede demostrar que en ambos casos la descomposición en factores primos es posible y única (el Teorema 12 es válido).

Definición. *Se llama dominio con factorización única [y también «dominio gaussiano (de Gauss)»] a un dominio de integridad en el cual*

1.° *Cualquier elemento no unidad puede descomponerse en factores primos;*

2.° *Esta descomposición factorial (o factorización) es única, salvo el orden y los factores unidad.*

Nuestro principal resultado va a ser que si G es un dominio gaussiano, también lo será cualquier dominio $G[x_1, \ldots, x_n]$ de formas polinómicas sobre G. Utilizando la inducción sobre n, se puede reducir el problema al caso $G[x]$, con una sola indeterminada, y éste es el caso que vamos a considerar.

En primer lugar, sumergiremos a G en su campo de cocientes F (Capítulo II, Teorema 7), y consideraremos $F[x]$ junto con $G[x]$. Típicamente, se puede imaginar a G como el dominio de los enteros, y a F como el de los racionales.

En segundo lugar, vamos a llamar «primitivo» a cualquier polinomio de $F[x]$ cuyos coeficientes, 1.°, pertenezcan a G (sean «en-

teros»), y 2.°, no tengan más divisores comunes que las unidades de G. De este modo, $3-5x^2$ es primitivo, $3-6x^2$ no lo es.

LEMA 1 (Gauss). *El producto de dos polinomios primitivos también es primitivo.*

Demostración. Pongamos $\sum_k c_k x^k = \sum_i a_i x^i \cdot \sum_j b_j x^j$. Si este producto no es primitivo, algún $p \in G$ dividirá a todas las c_k. Pero sean a_i y b_j los primeros coeficientes de $\sum_i a_i x^i$ y $\sum_j b_j x^j$, respectivamente, que no sean divisibles por p (los que existirán, ciertamente, si estos polinomios son primitivos). En tal caso, la fórmula (3) del coeficiente c_{i+j} en el producto da

$$a_i b_j = c_{i+j} - [a_0 b_{i+j} + \ldots + a_{i-1} b_{j+1} + a_{i+1} b_{j-1} + \ldots + a_{i+j} b_0],$$

así que el producto $a_i b_j$ es divisible por p, pues lo son todos los términos del segundo miembro. Esto significa que p debe figurar en la descomposición factorial única de uno al menos de los factores a_i o b_j, en contradicción con lo dicho al elegirlos.

LEMA 2. *Cualquier polinomio* $f(x)$ *en* $F[x]$, *distinto del cero, puede escribirse en la forma* $f(x) = c_f f^*(x)$, *con* c_f *en* F, *y siendo* $f^*(x)$ *primitivo. Además, para un* $f(x)$ *dado, la constante* c_f *y el polinomio primitivo* $f^*(x)$ *son únicos, salvo posibles factores unidades de* G.

Demostración. Pondremos $f(x) = (b_0/a_0) + (b_1/a_1)x + \ldots + (b_n/a_n)x^n$, a_i, $b_i \in G$ («enteros»). Con esto, si $c = 1/a_0 a_1 \ldots a_n$, tendremos $f(x) = cg(x)$, con los coeficientes de $g(x)$ perteneciendo a G. Ahora, sea c' el m. c. d. de los coeficientes de $g(x)$ (el cual existe, pues el teorema de factorización única es válido en G). Evidentemente, $f^*(x) = g(x)/c'$ es primitivo, y $f(x) = (cc')f^*(x)$. Este es lo que afirma la tesis, con $c_f = cc'$.

Para demostrar la unicidad de c_f y f^*, bastará ver que f^* es único, salvo unidades en G. A este fin, supongamos $f^*(x) = cg^*(x)$, donde $f^*(x)$ y $g^*(x)$ son primitivos y $c \in F$. Escribamos $c = u/v$, siendo u, $v \in G$ y primos entre sí, así que $ug^*(x) = vf^*(x)$. Los coeficientes de $ug^*(x)$ tendrán a v como factor común; luego, como u y v son primos entre sí, v dividirá a todos los coeficientes de $g^*(x)$. Pero $g^*(x)$ es primitivo, luego v es una unidad de G. Por simetría,

también u es unidad, y por ende lo será u/v. Esto completa la demostración.

A la constante c_f del Lema 2 la llamaremos factor *contingente* de $f(x)$; es única, salvo sus asociadas en G.

LEMA 3. *Si* $f(x)=g(x)h(x)$ *en* $G[x]$ *o también en* $F[x]$, *debe ser* $c_f \sim c_g c_h$ *y* $f^*(x) \sim g^*(x)h^*(x)$, *denotando por el signo «$\sim$» la relación de asociación en* $G[x]$.

Demostración. Por el Lema 1, el polinomio $g^*(x)h^*(x)$ es primitivo; también es igual, evidentemente, al producto de $f^*(x)$ por una constante; luego, por el Lema 2, ambos polinomios difieren sólo en un factor u unidad en G (son asociados); por tanto, $c_f=u^{-1}c_g c_h$.

Como corolario resulta que, si $f(x)$ pertenece a $G[x]$ y es reducible en $F[x]$, debe ser $f(x)=uc_f g^*(x)h^*(x)$. Esto proporciona la siguiente generalización del Teorema 1 del Capítulo III:

TEOREMA 13. *Un polinomio con coeficientes enteros, que puede descomponerse en factores polinomios con coeficientes racionales, puede también descomponerse en polinomios del mismo grado con coeficientes enteros.*

Es muy importante observar que, por el Lema 3, la factorización de cualquier $f(x)$ en $G[x]$ se divide en dos partes independientes: la descomposición del factor contingente c_f y la de su «parte primitiva» $f^*(x)$. La primera se desarrolla en G, y por hipótesis es posible y única. La segunda, por el Lema 3, equivale esencialmente a la factorización en $F[x]$, que es posible y única por el Teorema 12. Esto sugiere lo siguiente:

LEMA 4. *Si G es un dominio con factorización única, también lo es* $G[x]$.

Demostración. Por el Lema 2, cualquier polinomio $f(x)$ admite una factorización del tipo $f(x)=c_f f^*(x)$; luego para un elemento $f(x)$ primo en $G[x]$, uno de estos dos factores debe ser una unidad de $G[x]$. Por lo tanto, los elementos primos de $G[x]$ son de dos tipos: los primos p de G y los polinomios primitivos $q(x)$ que son irreducibles en $G[x]$ y también (Teorema 13) en $F[x]$.

Consideremos ahora cualquier polinomio $f(x)$ en $G[x]$. Éste admite una factorización en $F[x]$, y, por tanto, es asociado de un

producto de irreducibles primitivos de $G[x]$, como $f(x) \sim q_1(x)\ldots q_m(x)$. Por lo tanto, $f(x)=dq_1(x)\ldots q_m(x)$, pudiéndose descomponer el elemento d de G en factores irreducibles p_i de G. En resumen, $f(x)$ admite la descomposición

$$f(x)=p_1\ldots p_r q_1(x)\ldots q_m(x),$$

siendo cada p_i un primo de G, y cada $q_j(x)$ un polinomio irreducible primitivo de $G[x]$.

Los polinomios $q_j(x)$ que aparecen en esta descomposición factorial están unívocamente determinados, salvo unidades de G, como las partes primitivas de los factores irreducibles únicos de $f(x)$ en $F[x]$. Como las $q_j(x)$ son primitivas, el producto $p_1\ldots p_r$ es, esencialmente, el factor contingente único c_f de $f(x)$. Por lo tanto, las p_i son, esencialmente, los factores únicos de c_f en el dominio dado G. Esto demuestra que $G[x]$ es un dominio de factorización única.

Del Lema 4, por inducción sobre n, deducimos:

TEOREMA 14. *Si* G *es un dominio con descomposición factorial única, también lo es cualquier dominio* $G[x_1, \ldots, x_n]$ *de polinomios sobre* G.

En el Capítulo XIV, § 10, aparecerá un dominio que no es de factorización única, en el cual, por consiguiente, no son válidos ni el Teorema 10 ni el Teorema 12. (Cfr. el Ejercicio 14 g) del anterior § 5.)

EJERCICIOS

1. Representar como un producto de una constante por un polinomio primitivo de $J[x]$ los polinomios siguientes: $3x^2+6x+9$, $x^2/2+x/3+7$.
2. Presentar *todos* los divisores de $6x^2+3x-3$ en $J[x]$.
* 3. Describir un método sistemático para encontrar todos los factores lineales $ax+b$ de un polinomio $f(x)$ en $J[x]$.
4. ¿Para qué enteros n es $2x^2+nx-7$ reducible en $R[x]$?
5. Hallar los factores primos de los siguientes polinomios en $R[x]$:

$$x^3-1001x^2-1, \qquad x^4+50x^2+2.$$

6. Demostrar que dos elementos a y b en un dominio con factorización única, tienen siempre un m. c. d. (a, b) y un m. c. m. $[a, b]$.
7. Demostrar que $ab \sim (a, b)[a, b]$ en cualquier dominio con factorización única.

8. ¿Cuáles de las propiedades de los «primos relativos» establecidos en los Ejercicios 15 y 16 del § 7 son válidos en cualquier dominio de factorización única?
9. Con la notación del texto, demostrar directamente:
   a) Que $c_f f^*(x) \mid c_g g^*(x)$ en $G[x]$ si, y sólo si, $c_f \mid c_g$ en G y $f^*(x) \mid g^*(x)$ en $F[x]$;
   b) Mediante a), demostrar que un elemento «primo» de $G[x]$ que divida a un producto $a(x)b(x)$ debe dividir a uno de los factores.
10. Si $f(x)$ y $g(x)$ son primos entre sí en $F[x]$, demostrar que $yf(x)+g(x)$ es irreducible en $F[x, y]$.
11. Descomponer en factores irreducibles en $R^*[x, y]$ cada uno de los siguientes elementos; y probar que los factores obtenidos son irreducibles:

    a) x^3-y^3, b) x^4-y^2, c) x^6-y^6, d) $x^7+2x^3y+3x^3+9y$.

12. Hallar todos los polinomios irreducibles de grado 2 o menor en $J_2[x, y]$.
13. Demostrar que no existe en $R[x, y]$ ningún polinomio solución de la ecuación $1=s(x, y)(x-2)+t(x, y)(x+y-3)$.
14. Demostrar que el polinomio $f(x, y)$ es irreducible en $F[x, y]$ si hay una substitución $x \to t^r$, $y \to t^s$ que dé una forma polinómica $f(t^r, t^s)$ irreducible en $F[t]$, supuesto que el grado de $f(t^r, t^s)$ es el máximo de los enteros mr y ns, para todos los pares m, n de los exponentes que aparecen en algún término $x^m y^n$ de f.
*15. (Kronecker). Si $f(x) \mid g(x)$ en $J[x]$, demostrar que $f(c) \mid g(c)$ para cada c en J. Desarrollar a partir de aquí (y de la fórmula de interpolación de Ejercicio 3, § 2) un método sistemático para investigar con un número finito de operaciones todos los factores de un grado dado de cualquier $f(x)$ de $J[x]$.
16. Sea D el conjunto de todos los números racionales que pueden escribirse como fracciones a/b con un denominador b primo con 6. Demostrar que D es un dominio de factorización única.

* 9. Criterio de Eisenstein

La descomposición de un polinomio cualquiera en factores irreducibles es muy laboriosa. Existen procesos sistemáticos para llegar a establecer la descomposición de un polinomio dado, con coeficientes enteros, mediante un número finito de operaciones. En casos particulares es preferible, si puede ser, utilizar criterios especiales. Existen numerosos criterios de irreducibilidad, que dependen de las condiciones de divisibilidad de los coeficientes, y el más sencillo de ellos es el debido a Eisenstein:

TEOREMA 15. *Sea* $a(x)=a_n x^n+a_{n-1}x^{n-1}+\dots+a_0$ *un polinomio con coeficientes enteros, y sea* p *un número primo. Si es* $a_n \not\equiv 0$ *(módulo* p*),* $a_{n-1}\equiv a_{n-2}\equiv\dots\equiv a_0\equiv 0$ *(mód.* p*) y* $a_0 \not\equiv 0$ *(mód.* p^2*), el polinomio* $a(x)$ *será irreducible.*

Demostración. En cualquier posible factorización ($n=m+k$),

$$a(x)=(b_m x^m+b_{m-1}x^{m-1}+\ldots+b_0)(c_k x^k+c_{k-1}x^{k-1}+\ldots+c_0)$$

podemos suponer, por el Teorema 13, que ambos factores tienen coeficientes enteros, b_i y c_j. Como $a_0=b_0c_0$, la hipótesis $a_0\not\equiv 0$ (módulo p^2) implica que los factores b_0 y c_0 no son ambos divisibles por p. Para fijar las ideas, supongamos que $b_0\not\equiv 0$ (mód. p), mientras que $c_0\equiv 0$ (mód. p). Tomemos el más pequeño índice $r\leqslant k$ para el cual $c_{r-1}\not\equiv 0$ (mód. p), mientras que $c_{r-1}\equiv\ldots\equiv c_0\equiv 0$ (mód. p). En tal caso,

$$a_r=b_0c_r+b_1c_{r-1}+\ldots+b_rc_0\equiv b_0c_r \quad (\text{mód. } p).$$

Pero $b_0\not\equiv 0$ y $c_r\not\equiv 0$ dan $a_r\not\equiv 0$, puesto que p es primo. Por la hipótesis, el único coeficiente a_r con el que esto puede suceder es con el a_n, así que $r=n$; el grado del segundo de los supuestos factores debe ser n, así que el polinomio $f(x)$ es, efectivamente, irreducible.

Este criterio puede aplicarse al polinomio llamado «ciclotómico»

$$\phi(x)=(x^p-1)/(x-1)=x^{p-1}+x^{p-2}+\ldots+x+1 \tag{8}$$

(cuyas raíces son las raíces complejas p-ésimas de la unidad, que serán tratadas en el Cap. V). Ciertamente, el criterio de Eisenstein no se aplica inmediatamente a (8), pero un simple cambio de variable, $y=x-1$, salva el inconveniente, ya que el desarrollo del binomio da

$$(x^p-1)/(x-1)=[(y+1)^p-1]/y=$$
$$=y^{p-1}+py^{p-2}+\frac{p(p-1)}{1\cdot 2}y^{p-3}+\ldots+p.$$

Los coeficientes binómicos que aparecen a la derecha son todos enteros divisibles por p (primo), pues p aparece en cada numerador y no puede reducirse con los factores, más pequeños, del denominador. Como el polinomio en y satisface las condiciones para aplicar el criterio de Eisenstein, es irreducible ; de donde resulta la irreducibilidad del primitivo polinomio ciclotómico $\phi(x)$ de (8).

EJERCICIOS

1. ¿Cuáles de los siguientes polinomios son irreducibles sobre el campo racional?

x^3+2x^2+4x+2, $\quad x^3+2x^2+2x+4$, $\quad x^4-47$, $\quad x^4+15$.

2. Utilizar el criterio de Eisenstein y demostrar que x^2+1 es irreducible sobre R.
3. Si $f(x)$ es irreducible sobre un campo F, demostrar que $f(x+a)$ también lo es para cualquier a en F.
4. Si un polinomio $f(x)$ de grado $n > k$ satisface a la hipótesis $a_n \not\equiv 0$, $a_k \not\equiv 0$, $a_{k-1} \equiv \dots \equiv a_0 \equiv 0$ (mód. p) y $a_0 \not\equiv 0$ (mód. p^2), demostrar que $f(x)$ tiene un factor irreducible de grado k al menos.

* 5. Demostrar la irreducibilidad de un polinomio de grado impar $2n+1$ cuando se cumplen las condiciones: $a_{2n+1} \not\equiv 0$ (mód. p), $a_{2n} \equiv \dots \equiv a_{n+1} \equiv 0$ (mód. p), $a_n \equiv a_{n-1} \equiv \dots \equiv a_0 \equiv 0$ (mód. p^2), $a_0 \not\equiv 0$ (mód. p^3).

6. a) Si $f(x)$ es un polinomio mónico con coeficientes enteros, demostrar que la irreducibilidad de $f(x)$ módulo p implica su irreducibilidad sobre R.
   b) Demostrar que todos los factores de $f(x)$ sobre J se reducen módulo p a factores del mismo grado sobre J_p.
   c) Usar esto para probar, utilizando el menor p, la irreducibilidad sobre R de

$$x^3+6x^2+5x+25, \qquad x^3+6x^2+11x+8, \qquad x^4+8x^3+x^2+2x+5.$$

7. a) Sea $F[t]$ el dominio de todos los polinomios con una indeterminada t. Establecer y demostrar un criterio análogo al de Eisenstein para polinomios $f(x)$ con *coeficientes* en $F[t]$. (*Sugerencia*: Poner t en lugar de p.)
   b) Utilizar esto para demostrar la irreducibilidad de $x^5+3t^2x^3+2tx^2+t^4x+7t+t^3$ en el dominio $F[t, x]$.
8. a) Demostrar que $a+bx+cx^2$ es irreducible en $R^*[x]$ si, y sólo si, $b^2-4ac < 0$.
   b) Demostrar que $a+bx+cx^2+dx^3$ siempre es reducible en $R^*[x]$, para $d \neq 0$.

* 10. Fracciones simples

El teorema de unicidad en la descomposición factorial de los polinomios, puede aplicarse a las formas racionales, para obtener cierta representación simplificada de las mismas, semejante a la descomposición en fracciones simples utilizada en el Cálculo Integral.

Consideremos primero una forma racional $b(x)/a(x)$, en la que el denominador admite la factorización $a(x)=c(x)d(x)$, con los factores $c(x)$ y $d(x)$ primos entre sí. Por el Teorema 10, se tendrán dos polinomios $s(x)$ y $t(x)$ con $1=sc-td$; por tanto,

$$(9) \qquad f(x)/[c(x)d(x)] = [s(x)f(x)]/d(x) - [t(x)f(x)]/c(x).$$

Este resultado se enuncia así:

LEMA 1. *Una forma racional cuyo denominador es el producto de dos formas polinómicas c(x) y d(x) primas entre sí, puede expresarse como suma de dos cocientes con denominadores c(x) y d(x), respectivamente.*

Si el denominador $a(x)$ es una potencia, $a(x) = [c(x)]^m$, el proceso no se aplica directamente. Sin embargo, dividiendo el numerador por $c(x)$ como en el algoritmo de la división, resulta $b(x) = q_0(x)c(x) + r_0(x)$. Dividiendo de nuevo por $c(x)$ el cociente $q_0(x)$, obtenemos $q_0(x) = q_1c + r_1$. Combinando ambos resultados queda

$$b(x) = q_1(x)[c(x)]^2 + r_1(x)c(x) + r_0(x).$$

Repitiendo este proceso (esta frase encubre una inducción; quítesele el disfraz), resulta, en notación abreviada (*),

(10) $$b(x) = q_{m-1}c^m + r_{m-1}c^{m-1} + \dots + r_1c + r_0,$$

donde cada polinomio $r_i = r_i(x)$, si no es cero, tiene grado menor que $c(x)$. La forma racional $b(x)/a(x)$ resulta ahora

(11) $$b/c^m = q_{m-1} + r_{m-1}/c + r_{m-2}/c^2 + \dots + r_1/c^{m-1} + r_0/c^m.$$

Esto demuestra:

LEMA 2. *Una forma racional que tiene una potencia* $[c(x)]^m$ *como denominador, puede expresarse como un polinomio más la suma de formas racionales, cuyos denominadores son potencias de* c(x), *y cuyos numeradores tienen grado inferior al de* c(x).

Para combinar estos resultados, descompongamos un denominador arbitrario dado $a(x)$ en producto de polinomios mónicos irreducibles. Agrupando los factores iguales, resultará

(12) $$a(x) = a_0[p_1(x)]^{m_1}[p_2(x)]^{m_2}\dots[p_k(x)]^{m_k},$$

con exponentes enteros m_i. Dos polinomios mónicos irreducibles $p_1(x)$ y $p_2(x)$ son ciertamente primos entre sí, así que las potencias $[p_1(x)]^{m_1}$ y $[p_2(x)]^{m_2}$ también lo serán. Apliquemos ahora el

(*) Esto es exactamente análogo al desarrollo decimal de un entero, explicado en Capítulo I, § 12.

Lema 1 a la descomposición del denominador, en la que un factor es $c(x)=[p_i(x)]^{m_i}$, y el otro factor, $d(x)$, es todo lo que resta en la expresión (12). La repetición de esto da para f/g una suma de fracciones, cada una con denominador $[p_i(x)]^{m_i}$, a las cuales puede aplicarse la reducción expresada por (11).

TEOREMA 16. *Una forma racional* b(x)/a(x) *puede expresarse como un polinomio en* x *más una suma de fracciones («simples») de la forma* r(x)/[p(x)]^m, *siendo* p(x) *irreducible y* r(x) *de menor grado que* p(x). *Los denominadores* [p(x)]^m *que aparecen, son todos divisores de* a(x).

Si se desea explícitamente la descomposición de una fracción dada $b(x)/a(x)$, pueden recorrerse los distintos pasos de la demostración del Teorema 16, llegándose así al resultado pedido. Estas demostraciones, que pueden utilizarse para la obtención del objeto a que se refieren, son llamadas «constructivas». En algunas cuestiones matemáticas se establece la validez de ciertos teoremas de existencia, pero frecuentemente las demostraciones no son constructivas.

Consideremos, por ejemplo, sobre el campo de los números reales, $(x+1)/(x^3-1)$. El denominador es $(x-1)(x^2+x+1)$, y el segundo factor es irreducible. Por el algoritmo de la división se encuentra $x^2+x+1=(x+2)(x-1)+3$. Por lo tanto, multiplicando por el numerador $x+1$,

$$3(x+1)=(x+1)(x^2+x+1)-(x^2+3x+2)(x-1);$$

$$\frac{3(x+1)}{x^3-1}=\frac{x+1}{x-1}-\frac{x^2+3x+2}{x^2+x+1}.$$

Cada una de las fracciones que resultan puede simplificarse por una división posterior (*).

Sobre el campo R^* de los números reales, los únicos polinomios irreducibles son los lineales y los cuadráticos ax^2+bx+c, con $b^2-4ac<0$. (Este hecho se establece en Cap. V.) Por lo tanto, sobre R^*, cualquier fracción racional puede expresarse como suma de fracciones simples cuyos denominadores son potencias de expre-

(*) Compárese lo directo de este método con el que suele usarse en los textos de Cálculo, donde se deben hallar los «coeficientes indeterminados» A, B, C, que aparecen en los términos $A/(x-1)$, $(Bx+C)/(x^2+x+1)$.

siones lineales y cuadráticas. Esto se utiliza en Cálculo Integral para demostrar que la integral indefinida de cualquier función racional puede expresarse mediante funciones elementales (algebraicas, trigonométricas y exponencial, con sus inversas). Por el Teorema 16, la forma racional que debe integrarse es, esencialmente, una suma de términos de los tipos $c(x+a)^{-m}$ y $c(x+d)(x^2+ax+b)^{-n}$. Luego el anterior teorema sobre la integral se reduce a integrar estos dos tipos de funciones simples (lo cual puede hacerse).

EJERCICIOS

1. Descomponer en fracciones simples (sobre el campo real):

a) $\dfrac{3x+4}{x^2+3x+2}$, b) $\dfrac{1}{x^2-a^2}$, c) $\dfrac{1}{x^3+x}$,

d) $\dfrac{a^3}{x^3-a^3}$, e) $\dfrac{3}{x^4+5x^2+4}$, f) $\dfrac{3x-7}{(x-2)^2}$.

2. Descomponer $(4x+2)/(x^3+2x^2+4x+8)$:
   a) Sobre el campo J_5 de enteros mód. 5;
   b) Sobre el campo R de los números racionales.
3. a) Si un denominador es $a(x)=(x-r)(x-s)$, mientras que el numerador $b(x)$ tiene grado 1 o menos, demostrar que las fracciones simples tienen numeradores $b(r)/(r-s)$ y $b(s)/(s-r)$, supuesto $r\neq s$.
   b) Extender este resultado a los denominadores de grado n.
4. Demostrar la igualdad (10) por inducción sobre m.
5. Demostrar con detalle, por inducción, el Teorema 16.
6. a) Demostrar que cualquier forma racional no polinomia puede representarse como un polinomio más otra forma racional en la que el numerador es 0 o tiene grado inferior al denominador.
   b) ¿Es única esta representación?
7. Si se impone a todas las fracciones, incluso las simples, el tener un numerador de grado inferior al respectivo denominador, demostrar que
   a) La representación establecida en el Lema 1 es única;
   b) Lo mismo en el Lema 2;
   c) Lo mismo en el Teorema 16.
8. ¿Qué resulta de la afirmación de unicidad del Ejercicio 7 si el grado del numerador no está sometido a tal restricción?
9. a) Si $p(x)$ es irreducible, demostrar que en cualquier representación de una fracción $b(x)/p(x)$ (con b y p primos entre sí) como suma de fracciones, debe intervenir al menos una fracción con un denomina-

dor divisible por $p(x)$. (Esto significa que la ulterior descomposición en fracciones simples de $b(x)/p(x)$ está fuera de lugar.)

b) ¿Puede decirse lo mismo para $b(x)/[p(x)]^m$?

*10. Hallar la suma de

$$[(x+1)(x+2)]^{-1}+2[(x+2)(x+4)]^{-1}+\ldots+2^n[(x+2^n)(x+2^{n+1})]^{-1}.$$

*11. Desarrollar un método para representar cualquier número racional como una suma de «fracciones simples» de la forma especial a/p^n (p primo, $0 \leq a < p$). Por ejemplo. $1/6=1/2-1/3$.

CAPÍTULO V

Números complejos

1. Definición

El sistema de los números reales ha tenido que ampliarse con el de los números complejos, que ahora definiremos. Esta ampliación es de necesidad primordial en Álgebra, pero es también necesaria en otras ciencias matemáticas y físicas (como en Teoría de Funciones, Electromagnetismo, etc.). Ciñéndonos al Álgebra, demostraremos, además, que ésta es la ampliación natural del sistema de los números reales, si se pretende que toda ecuación polinómica tenga raíces.

DEFINICIÓN. *Un «número complejo» es un par* (x, y) *de números reales* x *e* y *a los que se llama, respectivamente, componentes real e imaginaria del complejo. Estos números se suman y se multiplican por las reglas*

(1) $$(x, y)+(x', y')=(x+x', y+y')$$

(2) $$(x, y)\cdot(x', y')=(xx'-yy', xy'+yx')$$

El sistema de los números complejos así definido se indicará por C.

La anterior definición no ha surgido por una revelación sobrenatural, sino por lo que vamos a explicar. Primero se observó que la ecuación $x^2=-1$ no tiene raíces reales (x^2 no puede ser negativo). Esto sugirió la introducción de un número imaginario i, cumpliendo la igualdad $i^2=-1$ y satisfaciendo, además, a las leyes or-

dinarias del cálculo algebraico. En lenguaje preciso, esta sugestión consiste en la hipótesis plausible de que exista un dominio de integridad D que contenga a un tal elemento i y a todo el dominio real R^*.

Cualquier expresión de la forma $x+yi$ (x e y números reales) será un elemento de D. Además, por la definición de dominio de integridad (leyes ordinarias del cálculo algebraico), se tendrá

(1′) $$(x+yi)\pm(x'+y'i)=(x\pm x')+(y\pm y')i,$$

(2′) $$(x+yi)\cdot(x'+y'i)=xx'+(xy'+yx')i+yy'i^2.$$

Como además $i^2=-1$, obtenemos de (2′)

(2″) $$(x+yi)(x'+y'i)=(xx'-yy')+(xy'+yx')i.$$

Como corolario resulta que el subdominio de D engendrado por R^* y por i contiene a todos los elementos de la forma $x+yi$, y sólo a ellos.

Por otra parte, $(x+yi)=(x'+y'i)$ implica $(x-x')=(y'-y)i$, y elevando al cuadrado ambos miembros, $(x-x')^2=-(y'-y)^2$, y como $(x-x')^2\geqslant 0$, $-(y'-y)^2\leqslant 0$, sólo es posible que $x-x'=y'-y=0$, o sea, $x=x'$, $y=y'$. En resumen, dos pares distintos de números reales (x, y), determinan elementos distintos $(x+yi)$ de D. Así se establece una correspondencia biunívoca $(x, y)\leftrightarrow x+yi$ entre los elementos de C y los del subdominio de D engendrado por R^* e i. Finalmente, comparando las fórmulas (1) y (2) con las (1′) y (2′) veremos que la correspondencia conserva sumas y productos, luego es un isomorfismo. Por lo tanto,

TEOREMA 1. *Sea D cualquier dominio de integridad que contenga al sistema* R* *de los números reales y a* i, *raíz cuadrada de* —1. *El subdominio de* D *engendrado por* R* *e* i *es isomorfo con* C.

Demostraremos ahora nuestra conjetura de que, efectivamente, existe un dominio de integridad D que contiene a los números reales y a una raíz cuadrada de —1.

TEOREMA 2. *El sistema de los números complejos definido anteriormente es un campo que contiene un subcampo isomorfo con* R* *y una raíz de* $x^2+1=0$.

Demostración. Por lo que hace a la adición, puesto que las componentes real e *imaginaria* se suman separadamente, es inmediato comprobar que las leyes conmutativa y asociativa son válidas, que (0, 0) es el elemento idéntico y que $(-x, -y)$ es el inverso aditivo (o sea, el opuesto) de (x, y).

Por lo que hace al producto, puede demostrarse que obedece a las leyes asociativa y conmutativa, que (1, 0) es el elemento idéntico, y que todo $(x, y) \neq (0, 0)$ tiene un inverso dado por

$$(3) \qquad (x, y)^{-1} = [x/(x^2+y^2), \; -y/(x^2+y^2)].$$

Todas estas afirmaciones se demuestran inmediatamente si se tiene en cuenta que, como veremos en § 2, el «argumento» y el «módulo» de los números complejos se combinan independientemente en la multiplicación, y de tal modo, que satisfacen a las leyes apuntadas. Pero en este momento es preferible comprobar estas leyes por sustitución directa en la definición (2). Sólo el cálculo para la ley asociativa es algo largo. Omitimos los detalles.

Finalmente, vamos a comprobar la ley distributiva, también por sustitución directa. Así pues, sean $z = (x, y)$, $z' = (x', y')$, $z'' = (x'', y'')$. Sustituyendo en (1)-(2) queda :

$$\begin{aligned} z(z'+z'') &= (x, y)\,(x'+x'', \; y'+y'') = \\ &= [x(x'+x'') - y(y'+y''), \; x(y'+y'') + y(x'+x'')] \, ; \\ zz'+zz'' &= (xx'-yy', \; xy'+yx') + (xx''-yy'', \; xy''+yx'') = \\ &= (xx'-yy'+xx''-yy'', \; xy'+yx'+xy''+yx'') \, ; \end{aligned}$$

luego $z(z'+z'') = zz'+zz''$, como queríamos demostrar

En este campo C de pares de números reales encontraremos un subcampo de números reales, utilizando la correspondencia $(x, y) \leftrightarrow$ $\leftrightarrow x+yi$ (ya empleada en la demostración del Teor. 1), en la cual los números reales se corresponden con los pares de segunda componente cero, y el par (0, 1) con i. Detallando más, si las segundas componentes, y e y', de las definiciones (1) y (2), son ambas nulas, se observa que las primeras componentes x y x', se suman y multiplican exactamente como los números reales. Esto es, precisamente, la prueba de que la correspondencia $x \leftrightarrow (x, 0)$ es un isomorfismo entre el campo R^* y un subcampo de C. Procediendo como en casos precedentes, cada número complejo $(x, 0)$ se identificará, simplemente, con el correspondiente número real x.

Finalmente, el par $(0, 1)$ es, como se presumirá, la deseada raíz de -1; en efecto, como caso particular de la definición (2) se demuestra que $(0, 1)^2=(-1, 0)=-1$. Por lo tanto, *definiremos* a i como par $(0, 1)$. Cualquier par (x, y) toma entonces la forma

$$(4)\qquad (x, y)=(x, 0)+(0, y)=(x, 0)+(y, 0)(0, 1)=x+yi.$$

La notación $x+yi$ se usará con preferencia a la (x, y), pues resulta más sugestiva. Por brevedad, escribiremos también $z=(x, y)=x+yi$, $w=(u, v)=u+vi$, $c=(a, b)=a+bi$, etc., es decir, que con una sola letra denotaremos un número complejo cuyas dos componentes sean las dos letras inmediatamente precedentes.

EJERCICIOS

1. Comprobar que la multiplicación de complejos es conmutativa y asociativa.
2. Comprobar que $(x, y)(x, y)^{-1}=(1, 0)$ cuando se utiliza la fórmula (3).
3. Resolver $(1, 1)(x, y)=(2, 1)$.
   a) Como un sistema de ecuaciones en x, y;
   b) Utilizando (3).
4. Hallar los complejos $z=x+iy$ y $w=u+vi$ satisfaciendo el sistema
   a) $z+iw=1$, $iz+w=1+i$;
   b) $(1+i)z-iw=3+i$, $(2+i)z+(2-i)w=2i$.
5. Hallar una raíz compleja de $x^2=-4$ y de $x^2=-a^2$, siendo a real.
6. Describir el subcampo de C engendrado por i y los números *racionales*.
7. ¿Es cierto el Teorema 1 cuando D es un anillo conmutativo con unidad? Dar detalles.
*8. Dar otra construcción de los números complejos, partiendo de una raíz ω de $x^2+x=-1$. Establecer un isomorfismo entre el campo resultante y C.
*9. Demostrar que si F es un campo ordenado, existe un campo F^* que contiene un subcampo isomorfo con F y una raíz cuadrada de -1.
*10. Utilizando los métodos de los teoremas 1 y 2, demostrar, sin recurrir a los números reales, que el campo racional R puede extenderse a un campo más amplio $R(\sqrt{2})$ que contiene a R y a una raíz cuadrada de 2.
*11. Demostrar que no es posible ninguna definición de «números complejos positivos» que convierta a C en un campo ordenado.

2. El plano complejo

Hay una representación biunívoca muy importante de los números complejos, sobre los puntos de un plano cartesiano. Cada número complejo $z=x+iy$ está representado por el punto $P=(x, y)$ cuya abcisa x es la componente real de z, mientras la ordenada y es su componente imaginaria.

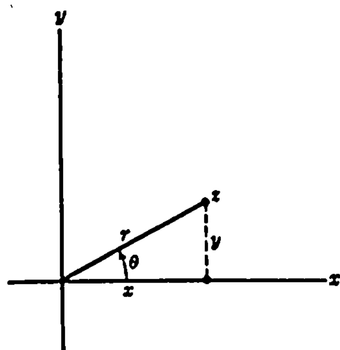


Figura 1

Las coordenadas polares se prestan también para esta representación. Recordemos que cada punto P del plano, y por ende cada número complejo z, está determinado unívocamente por dos coordenadas r y θ, siendo r la longitud (no negativa) del segmento $\overline{Oz}$ que une el punto P al origen, mientras que θ es el ángulo que forma este segmento con el eje de abcisas (fig. 1). Se llama *módulo* o *valor absoluto* de un complejo z, a la longitud r (no negativa), y *argumento* de z al ángulo θ. Por lo tanto,

(5) $$|z|=r=(x^2+y^2)^{1/2}, \qquad \arg z=\theta=\operatorname{arctg} y/x.$$

Estos dos valores determinan a x e y, pues

(6) $$x=r\cos\theta, \quad y=r\operatorname{sen}\theta, \quad z=r(\cos\theta+i\operatorname{sen}\theta),$$

según las fórmulas usuales para pasar de coordenadas polares a cartesianas.

La importancia del módulo y el argumento de un complejo es consecuencia de las fórmulas de Moivre, que pueden establecerse como sigue :

TEOREMA 3. *El módulo del producto de dos (o varios) números complejos es el producto de los módulos de los factores, y su argumento es la suma de los argumentos de los factores. En fórmulas,*

(7) $$|zz'|=|z|\cdot|z'|, \qquad \arg zz'=\arg z+\arg z'.$$

Demostración. Como vimos en (6), será $z=r(\cos\theta+i\operatorname{sen}\theta)$, $z'=r'(\cos\theta'+i\operatorname{sen}\theta')$. Sustituyendo en la definición (2) se obtiene,

$$zz'=rr'\left[(\cos\theta\cos\theta'-\operatorname{sen}\theta\operatorname{sen}\theta')+i(\cos\theta\operatorname{sen}\theta'+\operatorname{sen}\theta\cos\theta')\right];$$

y por fórmulas trigonométricas muy conocidas, esto da,

$$zz'=rr'\left[\cos(\theta+\theta')+i\operatorname{sen}(\theta+\theta')\right],$$

con lo que resulta (7).

Las desigualdades relativas al valor absoluto en la adición de números reales subsisten lo mismo para los números complejos. Es decir,

(8) $|z| > 0$ si no es $z=0$, $|0|=0$;

(9) $|z+z'| \leqslant |z|+|z'|$.

Para demostrarlo se observará que la fórmula (1) significa que la suma puede hallarse dibujando (figura 2) el paralelogramo del que tres vértices son z, 0 y z'. El cuarto vértice será $z+z'$. Las fórmulas (8) y (9) aparecen, pues, como conocidas desigualdades geométricas entre los valores absolutos de las longitudes de la figura.

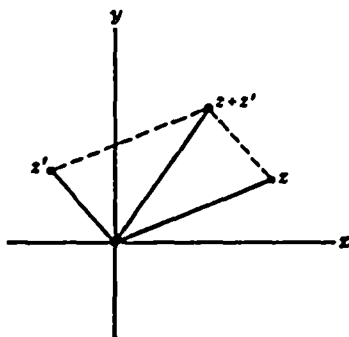


Figura 2

Procedamos ahora al cálculo de las raíces n-ésimas de la unidad. A partir de la fórmula (7) de Moivre, se ve inmediatamente que

$$[r(\cos\theta + i \operatorname{sen}\theta)]^{-1} = (1/r)[\cos(-\theta) + i \operatorname{sen}(-\theta)].$$

Además, vemos que $z^n=1$ si, y sólo si, $|z|^n=1$ y $n \cdot \arg z$ es un múltiplo entero $2k\pi$ de 2π. Como $|z| \geqslant 0$ y $\arg z$ es uniforme en el intervalo $0 \leqslant \theta < 2\pi$, existirán precisamente n soluciones de $z^n=1$. En coordenadas rectangulares, estas soluciones serán:

$$1;\ \cos\frac{2\pi}{n} + i \operatorname{sen}\frac{2\pi}{n};\ \ldots;\ \cos\frac{2(n-1)\pi}{n} + i \operatorname{sen}\frac{2(n-1)\pi}{n}$$

Si ponemos $\omega = \cos 2\pi/n + i \operatorname{sen} 2\pi/n$, otra representación de estas n raíces n-ésimas de la unidad será: 1, ω, ω^2, ..., ω^{n-1}. Geométricamente hablando, esto significa:

TEOREMA 4. *Las raíces n-ésimas complejas de la unidad son los vértices de un polígono regular de n lados inscrito en el círculo unidad* $|z|=1$.

Consideremos, más generalmente, la ecuación $z^n=c$, donde c es cualquier número complejo. En coordenadas polares, una solución es la siguiente:

$$z_0 = |c|^{1/n}(\cos\theta + i \operatorname{sen}\theta), \quad \text{con} \quad \theta = (1/n)\arg c.$$

Pero además wz_0 será también una raíz de $z^n=c$ si, y sólo si, $c=(wz_0)^n=w^nz_0^n=w^nc$, y por ende $w^n=1$. Así, las n raíces n-ésimas de c son: $z_0, \omega z_0, \omega^2 z_0, \ldots, \omega^{n-1}z_0$, donde ω es la definida antes. Es claro que estas raíces vienen también representadas por los vértices de un polígono regular.

Con la ayuda de las tablas logarítmicotrigonométricas es fácil el cálculo numérico de las raíces n-ésimas, $z_0, \omega z_0, \ldots, \omega^{n-1}z_0$ de $c=a+bi$. Partiendo de la identidad

$$\log |z_0|=\log |c|^{1/n}=(1/n) \log (a^2+b^2)^{1/2}=(1/2n) \log (a^2+b^2)$$

se puede calcular $|z_0|$. Por la fórmula (7), $\arg z_0=(1/n) \operatorname{arctg} (b/a)$ y $\arg \omega^k z_0=(1/n) \operatorname{arctg} (b/a)+360k/n$ (en grados sexagesimales). El cálculo se completa por la fórmula

$$z=r(\cos \theta+i \operatorname{sen} \theta)=|z| \cos (\arg z)+i\, |z| \operatorname{sen} (\arg z).$$

Cada raíz n-ésima compleja de la unidad, ω, satisface a una ecuación con coeficientes racionales, irreducible sobre el campo de los racionales. Estas ecuaciones son las llamadas «ciclotómicas», y juegan un importante papel en la teoría general de ecuaciones.

Por definición, cualquiera raíz n-ésima de la unidad satisface a la ecuación $z^n-1=0$; por lo tanto, todas ellas, excepto el 1, satisfacen a

$$(10) \qquad q_n(z)=(z^n-1)/(z-1)=z^{n-1}+z^{n-2}+\ldots+z+1.$$

En el Cap. IV, § 9, el criterio de Eisenstein permitió demostrar que $q_p(z)$ es irreducible si $n=p$ es primo.

Si n no es primo, las cosas no son tan simples. Así, si $n=4$, $z^3+z^2+z+1=(z+1)(z^2+1)$ es reducible. En general, aparecerá como factor de la (10) el polinomio ciclotómico satisfecho por las raíces k-ésimas de la unidad, siendo k divisor propio de n. Las raíces n-ésimas de 1, que no sean raíces de la unidad de orden $k<n$, se llaman *raíces primitivas* n-ésimas de la unidad (por ejemplo, las raíces primitivas cuartas de la unidad son i y $-i$). Todas ellas vienen expresadas por ω^m, siendo m primo con n, y todas satisfacen a una misma ecuación irreducible sobre el campo racional. Pero la demostración de estos resultados y el cálculo del grado de esta ecuación implican cuestiones de la teoría de números, que preferimos omitir aquí.

EJERCICIOS

1. Demostrar las leyes asociativa y conmutativa de la multiplicación, y la existencia de inversos, a partir de las fórmulas de Moivre.
2. Discusión de las raíces n-ésimas *reales* de la unidad. ¿Son racionales?
3. Calcular con 4 decimales las componentes real e imaginaria de las raíces cúbica y quinta de la unidad (utilizar tablas).
4. Calcular con 4 decimales las raíces cúbicas y cuartas de $2+2i$.
5. Enumerar las raíces primitivas de orden 12 de la unidad y hacer su representación sobre el círculo unidad.
6. Describir geométricamente la correspondencia $z \to zi$.
7. Hallar los factores irreducibles sobre R de $z^4 - 1$.
8. a) Demostrar que $\omega = \cos(2\pi/n) + i \operatorname{sen}(2\pi/n)$ es raíz primitiva n-ésima de 1.
   b) Demostrar que ω^m es raíz primitiva n-ésima de 1 si, y sólo si, n es primo con m.

*9. Se llaman *enteros de Gauss* a los números complejos $m+ni$, con m y n enteros. Demostrar:
   a) Los enteros de Gauss constituyen un dominio de integridad $J[i]$.
   b) Las únicas «unidades» de $J[i]$ son ± 1 y $\pm i$.
   c) Dados dos enteros de Gauss c y d, existen otros dos, q y r, tales, que $c = dq + r$, $|r| < |d|$.
   d) $J[i]$ es un dominio con descomposición factorial única.

3. Teorema fundamental del Álgebra

Vimos en § 1 que el sistema de los números complejos se obtiene adjuntando al sistema R^* de los números reales una raíz imaginaria i de la ecuación $z^2+1=0$. Pero ¿por qué nos detenemos aquí? ¿No intentaremos agregar ahora las raíces «imaginarias» de otros polinomios, para obtener así campos más amplios?

La respuesta a esta pregunta se encuentra en el llamado «Teorema fundamental del Álgebra» : pues, tan pronto como se adjunte i, cualquier ecuación polinómica tiene raíces complejas, así que no es necesario idear nuevos imaginarios para resolverla.

TEOREMA 5 (Euler-Gauss). *Cualquier polinomio* p(x) *de grado positivo con coeficientes complejos, tiene una raíz compleja.*

Se conocen muchas demostraciones de este célebre teorema. En todas ellas se implican conceptos no algebraicos, análogos a los introducidos en Cap. III. Hemos elegido una en que la parte no algebraica es de especial evidencia intuitiva.

Demostración. Como $p(z)=a_m z^m + a_{m-1}z^{m-1} + \ldots + a_0$, con $a_m \neq 0$, tiene las mismas raíces que

$$q(z)=z^m + (a_{m-1}/a_m)z^{m-1} + \ldots + (a_0/a_m) = z^m + c_{m-1}z^{m-1} + \ldots + c_0,$$

sólo será necesario considerar el caso en que el coeficiente principal sea la unidad.

Hecho esto, consideremos *dos* planos complejos, a los que llamaremos «plano z» y «plano w». La función dada $q(z)$ representa cada punto $z_0=(x_0, y_0)$ del plano z sobre un punto $w_0=q(z_0)$ del plano w Además, si z describe una línea continua del plano z, su imagen $q(z)$ describirá otra línea continua del plano w (ya que la función $q(z)$ es diferenciable). Nuestro objeto es demostrar que el punto 0 del plano w es la «imagen» $q(z)$ de algún punto del plano z; o, lo que es igual, que la representación de alguna curva del plano z pasa por el origen del plano w.

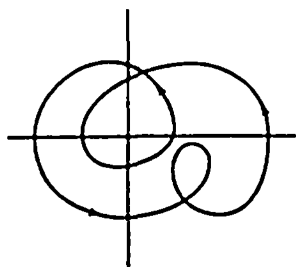


Figura 3

En particular, $q(z)$ representará a la circunferencia γ_r, de ecuación $|z|=r$, sobre otra curva cerrada γ'_r del plano w, como, por ejemplo, la dibujada en fig. 3. Consideremos ahora el número de veces, $n(r)$, que γ'_r *da una vuelta en torno al origen* en sentido directo (contrario a las agujas del reloj).

Es claro que si γ'_r no pasa por 0 (en cuyo caso el teorema estaría probado), $n(r)$ será un *entero* bien determinado (*).

Si $r=0$, γ'_r se reduce al solo punto $w=q(0)$. Veamos ahora que si r es bastante grande, $n(r)$ es el grado m de $q(z)$. En efecto, sea

$$q(z)=z^m + c_{m-1}z^{m-1} + \ldots + c_1 z + c_0 = z^m(1 + \sum_{k=1}^{m} c_{m-k}z^{-k}).$$

Por las fórmulas (7) de Moivre,

$$\arg q(z) = m \arg z + \arg (1 + \sum_{k=1}^{m} c_{m-k}z^{-k}).$$

Por lo tanto, como z describe el círculo γ_r en sentido directo, el cambio sufrido por arg $q(z)$ es la suma de m veces el cambio de

(*) Analíticamente se puede representar $n(r)$ por la integral curvilínea $(1/2\pi)\int d\theta$, a lo largo de γ_r' con $d\theta=(udv - vdu)/(u^2+v^2)$.

arg z (es decir, $2\pi m$), con el cambio de arg $(1+\sum_k c_{m-k}z^{-k})$. Pero si $|z|=r$ es suficientemente grande, por las fórmulas (8) y (9) $1+\sum_k c_{m-k}z^{-k}=w^*$ es interior al círculo $|w^*-1|<\frac{1}{2}$, luego no dará ningún giro en torno al origen (dibujar una figura que ilustre esto).

En conclusión : si r es lo bastante grande, $n(r)=m$: El incremento total de arg $q(z)$ es $2\pi m$. Pero si r cambia con continuidad, asimismo se deforma γ'_r (ya que $q(z)$ es función continua). Es geométricamente evidente (*) que si una curva que rodea al origen $n\neq 0$ veces se deforma hasta reducirse a un punto, deberá pasar por el origen en algún estado de su deformación. Por lo tanto, para algún r, γ'_r pasa por el origen ; pero esto dice, al fin, que para algún z es $q(z)=0$, c. q. d.

Como corolario se observará que si $p(z_1)=0$, por el teorema del resto (Cap. IV), podremos escribir $p(z)=(z-z_1)r(z)$. Si el grado m de $p(z)$ es mayor que 1, el cociente $r(z)$ tendrá también grado positivo, y por ende una raíz compleja $z=z_2$. Procediendo así, encontraremos m factores lineales de $p(z)$, o sea,

$$p(z)=c(z-z_1)(z-z_2)\dots(z-z_m). \tag{11}$$

De aquí se sigue que *los únicos polinomios irreducibles en el campo complejo son los lineales.* Una consecuencia de esto y del teorema de descomposición factorial única del Cap. IV es :

Teorema 6. *Cada polinomio con coeficientes complejos puede ser escrito de un modo, y sólo de uno, en la forma (11).*

Las raíces de $p(z)$ son, evidentemente, las z_i de (11), ya que el producto se anula si, y sólo si, uno de sus factores es nulo. Si un mismo factor $(z-z_i)$ aparece varias veces, el número de éstas es la *multiplicidad* de la raíz z_i. Mediante el cálculo se puede definir la multiplicidad de la raíz z_i como el «orden» con el cual $p(z_i)$ se anula en z_i, esto es : el mayor entero ν tal, que el polinomio $p(z)$ y sus $\nu-1$ primeras derivadas se anulan en z_i.

EJERCICIOS

1. Demostrar la unicidad de (11) sin utilizar el teorema general de unicidad del Cap. IV.

(*) En topología, esto es un teorema, cuya demostración puede verse en Alexandroff and H. Hopf, *Topologie*, Berlin, 1935, pág. 463.

* 2. Demostrar que cualquier «sucesión regular» de números complejos converge hacia un límite (cfr. Cap. III, § *6).
3. Demostrar que cualquier función racional compleja que sea finita para todo z es un polinomio.
4. Los pares (w, z) de números complejos, multiplicados y sumados por las reglas (1)-(2), ¿cuándo forman un anillo conmutativo con unidad? ¿Y un campo?
5. Demostrar que cualquier polinomio cuadrático puede llevarse a una de las formas $cz(z-1)$ o cz^2 por un automorfismo adecuado de $C[z]$.
6. a) Con el empleo de la serie de MacLaurin probar formalmente que $e^{ix}=\cos x+i \operatorname{sen} x$.
   b) Demostrar que cualquier número complejo puede escribirse como $re^{i\theta}$.
   c) Deducir las identidades $\cos z=(e^{iz}+e^{-iz})/2$, $\operatorname{sen} z=(e^{iz}-e^{-iz})/2i$.
* 7. Demostrar que cualquier función racional sobre el campo C de los números complejos puede escribirse como suma de un polinomio y de varias funciones con los numeradores constantes y cuyos denominadores son potencias de una función lineal (fracciones simples).
* 8. Demostrar que cualquier función racional puede integrarse mediante las funciones racionales y logarítmica compleja (es decir: inversa de la función exponencial).

4. Números conjugados y polinomios reales

En el sistema de los números complejos C, la ecuación $z^2=-1$ tiene las dos raíces i y $-i$, siendo $-i=0+(-1)i$. La correspondencia $x+yi \to x+y(-i)=x-yi$ transforma la primera de estas raíces en la segunda, e inversamente, mientras los números reales quedan inalterados. Además, esta correspondencia transforma sumas en sumas y productos en productos, como puede comprobarse por sustitución directa en las fórmulas (1) y (2), o por aplicación del Teorema 1. Dicho brevemente: la correspondencia es un *automorfismo* de C (un isomorfismo de C consigo mismo).

Podemos enunciar lo anterior más concisamente como sigue: Llamaremos *conjugado* z^* de un número complejo $z=x+yi$, al número $x-yi$. La correspondencia $z \to z^*$ es un automorfismo de período dos en C, puesto que

$$(12) \qquad (z_1+z_2)^*=z_1^*+z_2^*, \qquad (z_1z_2)^*=z_1^*z_2^*, \qquad (z^*)^*=z \quad (*)$$

Geométricamente, $z \to z^*$ equivale a la reflexión del plano complejo sobre el eje de abcisas; los únicos números iguales a sus conjugados son los reales.

(*) A la correspondencia $z \to z^*$ le llaman *conjugación* algunos autores. (N. del T.)

Los complejos conjugados intervienen en numerosas cuestiones de matemáticas y de física (especialmente en Mecánica Ondulatoria). Para su empleo, es conveniente tener en la memoria algunas fórmulas tan simples como

$$|z|^2 = zz^*, \qquad z^{-1} = z^*/|z|^2$$

Se emplean también para deducir la teoría de la descomposición factorial de polinomios en el campo real, que resulta fácilmente del

TEOREMA 7. *Un polinomio de coeficientes reales puede descomponerse en factores reales lineales y cuadráticos de discriminante negativo* (*).

Demostración. Sea $p(z)$ el polinomio dado; lo podemos escribir en la forma (11), donde las z_i son complejos (generalmente no reales). Como la correspondencia $z_i \to z_i^*$ aplicada a las raíces z_i es un automorfismo, hará corresponder a $p(z)$ otro polinomio $p^*(z) = c(z-z_1^*)(z-z_2^*)...(z-z_n^*)$, en el que cada coeficiente es el conjugado del correspondiente coeficiente de $p(z)$. Pero como los coeficientes de $p(z)$ son reales, $p(z) = p^*(z)$. Luego por la unicidad del desarrollo (11), $c = c^*$ es real, y las z_i serán también reales o complejas conjugadas a pares. Una z_i real dará el factor lineal $z - z_i$. Un par de raíces conjugadas $a+bi$, $a-bi$, con $b \neq 0$, pueden agruparse en

$$[z-(a+bi)][z-(a-bi)] = z^2 - 2az + (a^2+b^2)$$

dando así un factor cuadrático de $p(z)$ con discriminante $4a^2 - 4(a^2+b^2) = -4b^2 < 0$.

Recíprocamente, los polinomios lineales y de segundo grado con discriminante negativo son irreducibles sobre el campo *real* (estos últimos tienen sólo raíces complejas, y por ende carecen de factores lineales). Como corolario se desprende la unicidad de la descomposición factorial establecida en el Teorema 7.

Conviene destacar explícitamente que, como se ha dicho,

COROLARIO. *Las raíces complejas no reales de una ecuación polinómica con coeficientes reales aparecen a pares, cada una con su conjugada.*

(*) El discriminante de Ax^2+Bx+C es, por definición, $B^2 - 4AC$.

Esto generaliza al hecho bien sabido de que las dos raíces de la ecuación $ax^2+bx+c=0$ con discriminante $b^2-4ac<0$ son los dos complejos conjugados $x=(-b\pm\sqrt{b^2-4ac})/2a$.

EJERCICIOS

1. Resolver: a) $(1+i)z+3iz^*=2+i$,
   b) $zz^*+2z=3+i$,
   c) $zz^*+3(z-z^*)=4-3i$.
2. Resolver: $zz^*+3(z+z^*)=7$, $zz^*+3(z+z^*)=3i$.
3. Resolver el sistema

$$iz+(1+i)w=3+i, \quad (1+i)z^*-(6+i)w^*=4.$$

4. Dar una demostración independiente, del Cor. 2 del Teor. 3 del Cap. III.
5. Demostrar que si se adjunta el sistema de los números reales una raíz imaginaria de *cualquier* polinomio real irreducible no lineal, se obtiene un campo isomorfo con C.
6. Demostrar el Cor. del Teorema 7 sin utilizar la descomposición (11), observando que $z\to z^*$ es un automorfismo.
7. Demostrar que los automorfismos de C en que los números reales queden invariantes no pueden ser más que la identidad $(z\to z)$ y la conjugación $(z\to z^*)$.

* 5. Resolución de ecuaciones por radicales

En el § 3 se demostró teóricamente la existencia de las raíces de un polinomio dado con coeficientes complejos, pero no dimos ningún método para el cálculo efectivo de las mismas.

En esta sección daremos métodos para este cálculo, aplicables a las ecuaciones de grado no superior a *cuatro*.

El método consistirá en expresar la solución buscada, mediante una sucesión finita de operaciones con números fácilmente calculables. Cada una de estas operaciones será racional (adición, multiplicación, substracción y división) o bien la extracción de una raíz n-ésima. Las primeras operaciones han sido estudiadas en §§ 1 y 2; luego es evidente que nuestros resultados podrán aplicarse de modo general a cualquier campo en que pueden calcularse las raíces n-ésimas de sus elementos.

Consideramos primero el método de «completar el cuadrado» para resolver la ecuación cuadrática, conocido desde la enseñanza media. Tal ecuación,

$$ax^2+bx+c=0 \qquad (a\neq 0) \tag{13}$$

es equivalente (tiene las mismas raíces) que la ecuación sencilla

(14) $$x^2+Bx+C=0 \qquad (B=b/a,\ C=c/a).$$

Si ponemos $y=x+B/2$ (esto es, $x=y-B/2$), como para completar el cuadrado, vemos que (14) es equivalente a

(15) $$y^2=B^2/4-C.$$

Volviendo a poner x, a, b, c, en vez de y, B, C, esto da:

(16) $$x=y-B/2=(-b+\sqrt{b^2-4ac})/2a$$

teniéndose dos soluciones, según § 2.

La solución anterior era conocida por los matemáticos indios, y en forma geométrica por los griegos (cfr. Capítulo III, § 2). Las fórmulas análogas para resolver las ecuaciones de tercero y cuarto grado fueron descubiertas por dos matemáticos italianos del Renacimiento, Tartaglia (1530) y Ferrari (1545). Hasta el siglo XIX no fué demostrado por Abel y Galois la imposibilidad de resolver «por radicales» las ecuaciones de grado superior a cuatro (ver el Capítulo XV).

TEOREMA 8. *La ecuación cúbica general es resoluble por radicales.*

Demostración. Sea la ecuación $ax^3+bx^2+cx+d=0$; podemos dividir por a y reemplazar x por $y=x+b/3a$ (lo último es un tanteo previo para «completar el cubo»). La ecuación original se reduce entonces a la que sigue, en la cual falta el término con el cuadrado:

(17) $$y^3+py+q=0$$
$$(p=c/a-b^2/3a^2,\ q=d/a-bc/3a^2+2b^3/27a^3).$$

Sus raíces difieren de las de la ecuación original en $b/3a$.

Efectuemos ahora la sustitución de Vieta, $y=z-p/3z$. Reduciendo resulta:

(18) $$z^3-p^3/27z^3+q=0.$$

Multiplicando todo por z^3 resultará una ecuación de segundo grado en z^3, que resuelta por (16) nos dará:

(19) $$z^3=-q/2+\sqrt{q^2/4+p^3/27} \qquad \text{(dos valores).}$$

Esto da *seis* soluciones para z, en forma de raíces cúbicas. Sustituídas en la fórmula $y=z-p/3z$ quedan tres pares de soluciones para y, siendo iguales las dos soluciones de cada par. Así quedan tres valores para $x=y-b/3a$.

TEOREMA 9. *La ecuación cuártica general es resoluble por radicales.*

Demostración. Sea la ecuación completa $ax^4+bx^3+cx^2+dx+e=0$. Una vez más, dividiendo todo por a y reemplazando x por $z=x+b/4a$ obtenemos la ecuación

$$(20) \qquad z^4+pz^2+qz+r=0,$$

cuyas raíces difieren de las de la ecuación dada en $b/4a$. Pero, para cualquier u, (20) es equivalente a

$$(21) \qquad z^4+z^2u+u^2/4-z^2u-u^2/4+pz^2+qz+r=0, \qquad \text{o bien}$$

$$(z^2+u/2)^2-[(u-p)z^2-qz+(u^2/4-r)]=0.$$

El primer término es un cuadrado perfecto P^2, con $P=z^2+u/2$. El término entre corchetes será un cuadrado perfecto Q^2 si u es tal, que (igualando el discriminante a cero)

$$(22) \qquad q^2=4(u-p)(u^2/4-r);$$

esto da una ecuación cúbica a la que debe satisfacer u. Resolviendo esta ecuación cúbica según el Teorema 8 (¡y debe admitirse que el método es premioso!), y sustituyendo el valor adecuado de u, la ecuación cuártica puede escribirse en la forma $P^2-Q^2=$ $=(P+Q)(P-Q)$, o sea:

$$(23) \qquad (z^2+u/2+L)(z^2+u/2-L)=0,$$

donde L es una función lineal de z con coeficientes complicados. Sus raíces son, pues, las de los dos factores entre paréntesis, que pueden calcularse mediante (16).

En resumen, hallando una solución u de la cúbica (22), podemos reducir la ecuación cuártica dada a una «bicuadrada», esto es, producto de dos factores cuadráticos.

EJERCICIOS

1. Demostrar que para todo par y, p (complejos) existe un z satisfaciendo a $y=z-p/3z$. ¿Cuántos existen?
2. Resolver por radicales:

   a) $z^3+iz=2$,　　b) $z^3+3iz=1+i$,

   c) $z^3+3iz^2=10i$,　　d) $z^4-4z^3+(1+i)z=3i$.

3. Expresar una raíz de los Ejerc. 2 a) - 2 c), en forma decimal.
4. Demostrar, sin utilizar el teorema fundamental del álgebra, que todo polinomio real de grado $n<6$ tiene una raíz en el campo complejo.
5. a) Demostrar que la ecuación cúbica (17) con coeficientes complejos puede reducirse con la sustitución lineal $y=az$ a una de las dos formas $z^3+z=c$, $z^3=c$.

   b) Demostrar el teorema correspondiente para las ecuaciones cúbicas reales (con a real en la sustitución).

CAPITULO VI

Teoría de grupos

1. Simetrías del cuadrado

La noción de simetría es familiar a toda persona culta: las simetrías de una figura expresan las posibilidades de hacerla coincidir consigo mismo moviéndola sin deformarla. Pero pocos conocen que existe un álgebra consecuencia de la simetría. Esta álgebra es la que vamos a considerar ahora, en el caso concreto de las simetrías del cuadrado.

Imaginemos un cuadrado de cartulina, colocado sobre un plano en el que haya unos ejes cartesianos fijos, de modo que el centro del cuadrado caiga en el origen de coordenadas y uno de sus lados sea horizontal (paralelo al eje de abcisas). Claro es que el cuadrado tiene *simetría rotacional:* puede llevarse a coincidir consigo mismo por medio de los siguientes movimientos:

R : Rotación de 90°, en el sentido de las agujas de un reloj, alrededor de su centro.
R', R'' : Rotaciones análogas de 180° y 270°.

El cuadrado tiene también *simetría áxica:* esto quiere decir que coincide consigo mismo al reflejarse, en su plano, de los modos siguientes:

H : Reflexión en el eje horizontal que pasa por O.
V : Reflexión en el eje vertical que pasa por O.
D : Reflexión en la diagonal de los cuadrantes I-III.
D' : Reflexión en la diagonal de los cuadrantes II-IV.

Con esto, nuestra lista de simetrías alcanza a siete.

El álgebra de las simetrías tiene su origen en el hecho de que podemos *multiplicar* dos de estos movimientos, ejecutándolos sucesivamente. Así, el producto HR se obtiene por una reflexión previa en el eje horizontal, seguida de una rotación de 90° en el sentido de las agujas del reloj. Efectuando estos movimientos con la cartulina cuadrada se comprueba que su efecto total es el mismo que el de D', reflexión en la diagonal que une el vértice superior de la izquierda con el inferior de la derecha. La igualdad $HE = D'$ puede probarse de otro modo, viendo que ambos miembros producen el mismo efecto sobre cada vértice del cuadrado. Así, en la fig. 1, HR lleva primero 1 a 4, por H, y luego 4 a 3, por R; así que, en resumen, lleva 1 a 3, lo mismo que D'. De modo análogo, RH se define como el resultado de una rotación de 90° seguida de una reflexión en el eje horizontal. (*Nota:* El plano de la figura 1, que contiene a los ejes de reflexión, se considera inmóvil; los ejes no giran con el cuadrado.)

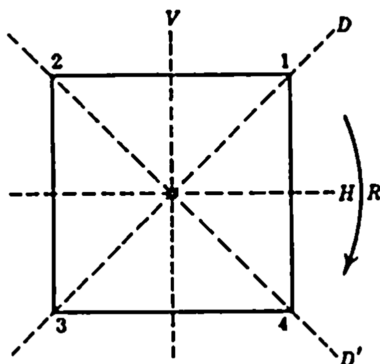


Figura 1

Una confrontación muestra que $RH = D \neq HR$, de lo que deducimos que nuestra «multiplicación» no es conmutativa en general. Sin embargo, es asociativa, como veremos en § 2.

El lector puede encontrar instructivo el calcular otros productos de simetrías del cuadrado (una lista completa se dará en el cuadro del § 4). Si lo hace así, descubrirá una excepción aparente a la regla según la cual, la sucesiva aplicación de dos simetrías cualesquiera es otra simetría. Si, por ejemplo, multiplica R por R' verá que su producto es un movimiento que deja fijos todos los puntos; se le llama «identidad» y se designa por I. De seguro que no se le considerará como una simetría por los no matemáticos; sin embargo, nosotros la consideraremos como una simetría (degenerada), significando así la posibilidad de multiplicar I con todas las simetrías ordinarias.

Si adoptamos este convenio, concluiremos que el «grupo» de simetrías del cuadrado tiene exactamente ocho elementos.

No sólo el cuadrado, sino cualquier polígono o poliedro regular (p. e., el cubo o el icosaedro regular) tienen grupos de simetrías que pueden hallarse siguiendo el método elemental que hemos iniciado antes.

También en muchas ornamentaciones aparecen simetrías interesantes. Consideremos el modelo sencillo indefinido



en el cual las flechas están espaciadas uniformemente de centímetro en centímetro a lo largo de una recta. Las tres simetrías elementales que aparecen son : T, traslación a la derecha en 1 cm. ; T', traslación a la izquierda en 1 cm., y H, reflexión en el eje horizontal. Todas las otras pueden obtenerse por multiplicación de las tres señaladas.

EJERCICIOS

1. Calcular HV, HD', $D'H$, $R'D'$, $D'R'$, $R'R'$
2. Describir TH y HT en el modelo ornamental considerado al final del párrafo anterior.
3. Enumerar las simetrías de un triángulo equilátero y calcular cinco productos.
4. Enumerar las simetrías de un rectángulo general y calcular todos sus productos.
* 5. ¿Cuántas simetrías posee el tetraedro regular? ¿Y el octaedro regular? Dibujar las figuras.
* 6. Demostrar que todas las simetrías del modelo ornamental del texto pueden obtenerse por multiplicaciones repetidas de H, T y T'.

2. Grupos de transformaciones (*)

Los principios que han presidido la precedente discusión, sobre el grupo de simetrías del cuadrado, pueden formularse con mayor generalidad.

Puede reemplazarse el cuadrado por un conjunto S de elementos cualesquiera, y en vez de «simetrías» se pueden considerar «transformaciones» ϕ de naturaleza muy general. Por ejemplo, S puede componerse de los números enteros positivos y ϕ puede ser la transformación que transporta cada número n sobre el que le sigue, $n+1$.

(*) Véase, además, Cap. XVI, § 2.

Por *transformación* uniforme ϕ de S en sí mismo se entiende una regla que asigna a cada elemento $p \in S$ un solo elemento imagen o transformado, el cual es, asimismo, un elemento de S, que se denota por $p\phi$ ó $\phi(p)$. Por brevedad, para referirnos a una transformación uniforme de S en sí mismo, diremos una «transformación en S»: éstas generalizan la noción de «movimiento rígido» del cuadrado, considerada en el § 1. Sin pérdida de generalidad podemos imaginarnos siempre al conjunto S como un «espacio» (p. ej., plano o esfera), y llamar «puntos» a los elementos de S.

El producto $\phi\phi'$ de dos transformaciones ϕ y ϕ' en S puede definirse como el resultado de ejecutarlas sucesivamente; primero ϕ y después ϕ'.

(1) $$p(\phi\phi') = (p\phi)\phi' \qquad \text{para todo } p,$$

la cual define el efecto del producto $\phi\phi'$ sobre cualquier elemento p de S.

La multiplicación de transformaciones verifica la

Ley asociativa: $\phi(\phi'\phi'') = (\phi\phi')\phi''$ para todo ϕ, ϕ', ϕ''.

Esto es fácil de ver intuitivamente, pues lo mismo $\phi(\phi'\phi'')$ que $(\phi\phi')\phi''$ suponen la ejecución primero de ϕ, luego de ϕ' y después de ϕ'', en este orden. Para probarlo formalmente hay que comenzar por definir la igualdad entre transformaciones: esto se hace diciendo que dos transformaciones son iguales cuando producen el mismo efecto sobre cada punto de S. Simbólicamente,

$$\psi = \psi' \quad \text{significa que} \quad p\psi = p\psi' \quad \text{para todo } p.$$

Con arreglo a tal definición, puede probarse que, para todo p.

$$p[\underset{\phi(\phi'\phi'')}{\phi(\phi'\phi'')}] = (p\phi)(\phi'\phi'') \underset{\phi'\phi''}{=} [(p\phi)\phi']\phi'' \underset{\phi\phi'}{=} [p(\phi\phi')]\phi'' \underset{(\phi\phi')\phi''}{=} p[(\phi\phi')\phi''].$$

pues cada paso supone la aplicación de la definición (1) de multiplicación al producto indicado debajo de cada igualdad. Esto prueba que $\phi(\phi'\phi'') = (\phi\phi')\phi''$.

La transformación idéntica I en el espacio S puede definirse como la transformación que deja fijos a todos los puntos. Esto puede establecerse algebraicamente por la identidad

(2) $$pI = p \qquad \text{para todo } p.$$

De las definiciones (1) y (2) sigue directamente la

Ley de identidad: $I\phi = \phi I = \phi$ para todo ϕ.

Para ver esto, basta notar que $p(I\phi) = (pI)\phi = p\phi$ para todo p, y análogamente que $p(\phi I) = (p\phi)I = p\phi$.

Las simetrías del cuadrado no son solamente uniformes, sino que son *biunívocas* (*). Esto quiere decir que cada punto q de S es el transformado de otro p, y sólo de éste. De aquí que la operación inversa ϕ^{-1}, que hace corresponder a q el elemento p del cual procedía, es también una transformación uniforme en S, si ϕ es biunívoca. En otras palabras, la inversa de una transformación biunívoca ϕ es la transformación ϕ^{-1} definida como sigue:

(3) $q\phi^{-1} = p$, si $p\phi = q$, y sólo en este caso.

Esta condición puede formularse mediante un producto: la sustitución de la segunda ecuación en la primera, y viceversa, da

$$p = q\phi^{-1} = (p\phi)\phi^{-1} = p(\phi\phi^{-1}), \qquad q = p\phi = (q\phi^{-1})\phi = q(\phi^{-1}\phi).$$

Esto nos enseña que cada producto $\phi\phi^{-1}$ y $\phi^{-1}\phi$ es igual a la transformación idéntica I; de aquí que para cualquier ϕ biunívoca tendremos:

Ley de inversa: $\phi\phi^{-1} = \phi^{-1}\phi = I$.

Recíprocamente, supongamos que una transformación ψ en S es una «inversa» de una dada ϕ en el sentido «formal» de esta ley, así que $\phi\psi = \psi\phi = I$. Estas ecuaciones nos dicen que, para cualesquiera p y q,

$$(p\phi)\psi = pI = p, \qquad (q\psi)\phi = qI = q.$$

La segunda igualdad afirma que cada punto q es el transformado $r\phi$ de algún punto $r = q\psi$. La primera dice que si q es el transformado $q = p\phi$ de algún punto p, entonces $q\psi = p$; así que q es el transformado de un único punto $r = q\psi$. Esto significa que ϕ es biunívoca. En otras palabras: hemos visto que $p\phi = q$ si, y sólo si, $p = q\psi$. Esto es afirmar que ψ satisface a la definición (3) de inversa

(*) Esto no es casual. Todas las transformaciones de la teoría de grupos son biunívocas. Sin embargo, las transformaciones sin inversa desempeñarán un importante papel cuando estudiemos las matrices (Cap. VIII).

de ϕ. En resumen, cualquier transformación ϕ con una ψ tal que $\phi\psi=\psi\phi=I$, es biunívoca, y se tiene $\phi^{-1}=\psi$. Por lo tanto, resulta :

TEOREMA 1. *Para que una transformación sea biunívoca es necesario y suficiente que tenga una inversa formal; esta inversa formal será una inversa en el sentido de la definición (3).*

En estas condiciones se dice que ϕ es una transformación *sobre* S (cfr. Cap. XVI, § 2). Estamos ahora en condiciones de definir el concepto de «grupo» de transformaciones. Por *grupo de transformaciones* en un espacio S se entiende un conjunto de transformaciones biunívocas en S y tales, que entre ellas está la identidad, si está una transformación está su inversa, y si están dos transformaciones está el producto de ambas.

TEOREMA 2. *El conjunto de todas las transformaciones biunívocas sobre un espacio cualquiera, es un grupo.*

Deberemos probar que la identidad I es biunívoca, que la inversa de una transformación biunívoca es biunívoca y que el producto de dos transformaciones biunívocas es otra biunívoca. Pero la ley de identidad da $II=I$, así que I es su propia inversa (formal) y es biunívoca por el Teorema 1. Además, si ϕ es biunívoca, la ley $\phi\phi^{-1}=\phi^{-1}\phi=I$ significa que ϕ^{-1} tiene ϕ como inversa formal; luego ϕ^{-1} es biunívoca, y

$$(\phi^{-1})^{-1}=\phi. \tag{4}$$

Finalmente, el producto de dos transformaciones biunívocas cualesquiera ϕ y ψ tiene una inversa, pues, por hipótesis,

$$(\phi\psi)(\psi^{-1}\phi^{-1})=\phi(\psi\psi^{-1})\phi^{-1}=\phi I\phi^{-1}=\phi\phi^{-1}=I,$$
$$(\psi^{-1}\phi^{-1})(\phi\psi)=\psi^{-1}(\phi^{-1}\phi)\psi=\psi^{-1}I\psi=\psi^{-1}\psi=I,$$

así, $\phi\psi$ es biunívoca y tiene una inversa,

$$(\phi\psi)^{-1}=\psi^{-1}\phi^{-1}. \tag{5}$$

El enunciado es: la inversa de un producto es el producto de las inversas, tomadas en orden contrario.

EJERCICIOS

1. Calcular VD, $(VD)R'$, DR', $V(DR')$ en el grupo del cuadrado.
2. Calcular, análogamente, HR, $R'(HR)$, $R'H$, $(R'H)R$.

3. Sea S el conjunto de todos los números reales (o de los puntos x de una recta); las transformaciones que vamos a considerar tienen la forma $x\phi = ax + b$. En cada uno de los siguientes casos, hallar cuándo el conjunto de todas las ϕ posibles, con coeficientes a y b del tipo indicado, constituye un grupo de transformaciones:

a) a y b números racionales.
b) $a=1$, b un entero impar.
c) $a=1$, b un entero positivo o nulo.
d) $a=1$, b un entero par.
e) a un entero, $b=0$.
f) $a \neq 0$, a y b números reales.
g) $a \neq 0$, a un entero, b un número real.
h) $a \neq 0$, a un número real, b un entero.
i) $a \neq 0$, a un entero, b un número irracional.
j) $a \neq 0$, a racional, b número real.

¿En cuáles de estos grupos es conmutativa la «multiplicación»?

4. Hallar todas las transformaciones sobre un «espacio» S de, exactamente, tres «puntos». ¿Cuántas hay? ¿Cuántas de ellas son biunívocas?
5. Demostrar que la correspondencia $n \to n^2$ sobre los enteros positivos no tiene inversa.
6. Calcular $[R^{-1}(VR)]^{-1}[(R^{-1}D)R]$ para el grupo del cuadrado.
7. Resolver la ecuación $RXR' = D$ para el grupo del cuadrado.
8. Hallar la inversa de cualquier simetría del cuadrado y comprobar la regla (4) en este caso.
9. Hallar la inversa de cualquier simetría del rectángulo y comprobar la regla (5).
10. Si $\phi_1, \ldots, \phi_n$ son biunívocas, demostrar que también lo es $\phi_1\phi_2 \cdots \phi_n$, con $(\phi_1\phi_2 \cdots \phi_n)^{-1} = \phi_n^{-1}\phi_{n-1}^{-1} \cdots \phi_1^{-1}$.

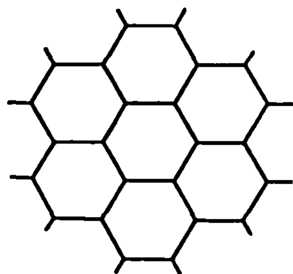
3. Ejemplos

En el estudio de la Geometría, constantemente surgen grupos de transformaciones sobre «espacios» variados. Muchos de estos grupos consisten, sencillamente, en las simetrías de tales espacios respecto a propiedades convenientemente elegidas.

Un ejemplo inmediato lo proporciona la consideración de las simetrías del cubo. Geométricamente hablando, éstas son las transformaciones biunívocas que conservan las distancias sobre el cubo. Reciben el nombre de «isometrías» y son 48 en total. Para ver esto, observemos que cualquier vértice inicial puede llevarse a coincidir con uno cualquiera de los ocho vértices. Después de fijar este primer vértice, los tres adyacentes pueden permutarse de seis modos, o sea que tenemos $6 \times 8 = 48$ posibilidades. Cuando un vértice y los tres adyacentes ocupan posiciones conocidas, cada punto del cubo

está en una posición fijada, así que la simetría está enteramente determinada. De aquí que el cubo tiene exactamente 48 simetrías. Muchas de ellas tienen propiedades geométricas especiales, tales como transportar cada punto sobre su diametralmente opuesto.

Un grupo conocido que contiene una infinidad de transformaciones es el llamado grupo euclídeo, formado por las «isometrías» del plano, o sea, en el lenguaje de la Geometría elemental, por las transformaciones que hacen al plano congruente consigo mismo. Está constituído por productos de traslaciones, rotaciones y reflexiones (simetrías áxicas); lo estudiaremos con mayor detalle en el capítulo IX.


Figura 2

Otro grupo es el de las transformaciones de semejanza del espacio, que son aquellas transformaciones biunívocas que multiplican todas las distancias por un factor constante $k > 0$ (factor de proporcionalidad). También constituyen grupo los movimientos de rotación de una superficie esférica sobre sí misma. Asimismo, las isometrías del plano que dejan invariante un reticulado de hexágonos regulares forman otro grupo (fig. 2). Para terminar: una cinta de goma, manteniendo sus extremos en dos puntos fijos P y Q, puede deformarse de muchos modos a lo largo del segmento PQ; todas estas transformaciones forman un grupo, llamado grupo de los «homeomorfismos» del segmento PQ.

En términos generales, aquellas transformaciones biunívocas de un conjunto de elementos, que conservan algunas propiedades de estos elementos, forman un grupo. Félix Klein, en su célebre *Programa de Erlangen* (1872), ha explicado brillantemente cómo las diferentes ramas de la Geometría consisten en el estudio de las propiedades de espacios convenientes, que se conservan en grupos de transformaciones apropiados. Volvemos sobre este tema en el Capítulo IX («Grupos lineales» y § 12).

EJERCICIOS

1. Describir todas las simetrías de una rueda con seis radios igualmente espaciados.
2. Describir las seis simetrías de un cubo con uno de sus vértices fijo. (*Sugerencia:* ¿Qué puede ocurrir con los otros tres vértices, adyacentes?)

3. Sean S, T las reflexiones de un cubo en planos paralelos a caras distintas. Describir geométricamente ST.
4. Describir algunas isometrías del plano que lleven sobre sí mismo el reticulado hexagonal de la figura 2.
5. Hacer lo mismo para un reticulado de cuadrados. ¿Pueden enumerarse todas estas transformaciones (esto es difícil)?
6. Lo mismo para una red de triángulos equiláteros, y relacionar esto con el grupo del Ejerc. 1.
7. Hacer lo mismo para un cilindro infinito, para un cilindro limitado, para una hélice que se arrolla al cilindro formando ángulo constante con el eje del mismo.
* 8. Demostrar que las transformaciones $x \rightarrow x' = (ax+b)/(cx+d)$ con $ad - bc = 1$ y con coeficientes de cualquier campo F, constituyen un grupo sobre F.

4. Grupos abstractos

No son los grupos de transformaciones los únicos sistemas que poseen una multiplicación que satisface a las leyes asociativa, idéntica y de inversa del § 2. Por ejemplo, los números distintos de cero de un campo (por ej., el racional, el real o el complejo) las satisfacen. El producto de dos números distintos de cero es un número diferente de cero (Cap. II, § 1) ; la ley asociativa es válida ; la unidad 1 del cuerpo satisface a la ley de identidad, y $1/x = x^{-1}$ proporciona el elemento inverso de x.

Análogamente, los elementos de cualquier dominio de integridad (esta vez incluyendo al cero) satisfacen a dichas leyes cuando se combinan por adición. Así, dos elementos tienen una suma determinada, la adición es asociativa, el cero satisface a la ley de identidad y $-x$ a la ley de inversa, relativa a la adición.

Es conveniente introducir el concepto de grupo abstracto para incluir en él estos casos y otros análogos.

DEFINICIÓN. *Un* grupo abstracto G *es un sistema de elementos cerrado para una operación binaria uniforme y asociativa ; además, con relación a esta operación,* G *contiene un elemento (llamado identidad o unidad) que satisface a la ley idéntica, y para cada elemento en* G *se encuentra otro (llamado su inverso) satisfaciendo a la ley de inversa.*

Al tratar de grupos abstractos denotaremos sus elementos por minúsculas latinas, a, b, c, ... La notación de producto «ab» será

npleada ordinariamente para indicar el resultado de aplicar la)eración propia del grupo a dos de sus elementos a y b ; sin em- ırgo, otras notaciones, tales como «$a+b$» y «$a \circ b$», son igual- ente válidas. Con la notación del producto, y poniendo «e» para elemento idéntico o unidad, las tres leyes que definen a los gru-)s abstractos se expresan así:

ey asociativa:	$a(bc)=(ab)c$,	para todo a, b, c.
ey idéntica:	$ae=ea=a$,	para todo a.
ey de inversa:	$aa^{-1}=a^{-1}a=e$.	para cada a y algún a^{-1}.

Si la operación del grupo satisface a la ley conmutativa, el :upo se llama *conmutativo* o grupo *abeliano*. Usando este con- ;pto, podemos dar la definición de campo en forma más elegante :

DEFINICIÓN. *Un campo es un sistema de elementos, cerrado ıra dos operaciones binarias, adición y multiplicación, tales que,*) *respecto a la adición,* F *es un grupo conmutativo con el cero* ımo *elemento idéntico; 2) respecto a la multiplicación, los ele- entos de* F *(sin el cero) forman otro grupo conmutativo; 3) la ultiplicación es distributiva con la adición.*

Estos postulados incluyen todos los establecidos antes para ca- ıcterizar un campo, excepto las leyes asociativa y conmutativa ıra productos con un factor cero; éstas pueden ser comprobadas, ;ro omitimos el detalle.

Algunos de los resultados de las primeras secciones de los Capí- ılos I-II resultan ahora como corolarios del siguiente teorema so- 'e grupos abstractos :

TEOREMA 3. *En todo grupo, las ecuaciones* $xa=b$ *y* $ay=b$ *tie- en las respectivas soluciones únicas* $x=ba^{-1}$ *e* $y=a^{-1}b$. *De aquí ıe* $ca=da$ *implique* $c=d$, *y lo mismo suceda con* $ac=ad$ *(ley de mplificación).*

Demostración. Evidentemente, $(ba^{-1})a=b(a^{-1}a)=be=b$, y aná- gamente $a(a^{-1}b)=b$. Recíprocamente, $xa=b$ implica $x=xe=xaa^{-1}$ ba^{-1}, y del mismo modo, $ay=b$ implica $y=a^{-1}b$, como queríamos ;mostrar.

Puesto que en cualquier grupo G, las ecuaciones $ex=e$, $ay=e$ tienen, por el Teorema 3, las soluciones únicas $x=e$, $y=a^{-1}$, resulta :

COROLARIO. *Un grupo abstracto tiene solamente un elemento unidad; y sólo un inverso* a^{-1} *para cada elemento* a.

TEOREMA 4. *En la precedente definición de grupo, las leyes de identidad e inversa pueden reemplazarse por las leyes más débiles:*

Unidad a la izquierda : *Para algún* e, $ea=a$ *para todo* a.
Inverso a la izquierda : *Dado* a, $a^{-1}a=e$ *para algún* a^{-1}.

Demostración. Si se cumplen estas leyes más débiles, la simplificación a la izquierda es posible, esto es, $ca=cb$ implica $a=b$. Para verlo, basta premultiplicar los dos miembros de $ca=cb$ por c^{-1} y aplicar la ley asociativa, obteniendo $(c^{-1}c)a=(c^{-1}c)b$, lo cual da $ea=eb$, y de aquí $a=b$. La unidad a la izquierda es también la unidad a la derecha, ya que

$$a^{-1}ae=ee=e=a^{-1}a,$$

de la que, por simplificación a la izquierda, $ae=a$ para todo a. Finalmente, los inversos por la izquierda son también inversos por la derecha, pues

$$a^{-1}(aa^{-1})=(a^{-1}a)a^{-1}=ea^{-1}=a^{-1}=a^{-1}e,$$

ya que la unidad por la izquierda es unidad por la derecha. La simplificación por la izquierda da $aa^{-1}=e$. Esto completa la demostración.

Hay muchos otros sistemas de postulados para caracterizar a los grupos. Uno muy útil puede establecerse, mediante la posibilidad de la división, así :

TEOREMA 5. *Si* G *es un sistema no vacío con una multiplicación asociativa, respecto a la cual todas las ecuaciones* $xa=b$ *y* $ay=b$ *tienen soluciones* x *e* y *en* G, *entonces* G *es un grupo.*

Dejamos la demostración como ejercicio para el lector (Ejercicio 12).

Utilizando sistemáticamente las leyes a que obedece la multiplicación en un grupo G, podemos construir una tabla de multipli-

car para conocer el producto de dos elementos de G, con tal que el número de elementos de G sea finito. Es un cuadro de doble entrada encabezado por los elementos del grupo, que también forman la primera columna. El elemento que pertenece a la fila y columna que respectivamente comienzan con a y b, es el producto ab (en este orden).

En la fig. 3 aparece la tabla de multiplicar para el grupo de las simetrías del cuadrado. El cálculo se hace repitiendo los que hicimos en § 1 para demostrar que $HR=D'$ y que $RH=D$.

	I	R	R'	R''	H	V	D	D'
I	I	R	R'	R''	H	V	D	D'
R	R	R'	R''	I	D	D'	V	H
R'	R'	R''	I	R	V	H	D'	D
R''	R''	I	R	R'	D'	D	H	V
H	H	D'	V	D	I	R'	R''	R
V	V	D	H	D'	R'	I	R	R''
D	D	H	D'	V	R	R''	I	R'
D'	D'	V	D	H	R''	R	R'	I

Figura 3

Muchas de las propiedades del grupo pueden verse directamente fijándose en su tabla de multiplicar. Así, la existencia de unidad indica que hay alguna fila y alguna columna que son repetición de las que respectivamente encabezan el cuadro. La posibilidad de resolver la ecuación $ay=b$ indica que en la fila a debe aparecer el elemento b; como la solución es única, b ha de estar una sola vez en esta fila. Para que el grupo sea conmutativo es necesario y suficiente que sea simétrico con relación a la diagonal principal. En cuanto a la propiedad asociativa, desgraciadamente no puede verse con facilidad en la tabla.

EJERCICIOS

1. Sean a, b, c, elementos fijos de un grupo. Demostrar que la ecuación $xaxba=xbc$ tiene una solución y sólo una.
2. En un grupo con $2n$ elementos, demostrar que existe un elemento distinto de la identidad el cual es su propio inverso.
3. ¿Forman grupo los números reales positivos con la adición? ¿Y con la multiplicación? ¿Lo forman los enteros pares con la adición? ¿Y los impares? ¿Por qué?
4. En el campo J_{11} de enteros módulo 11, ¿cuáles de los siguientes conjuntos son grupos para la multiplicación?
   a) (1, 3, 4, 5, 9); b) (1, 3, 5, 7, 8); c) (1, 8); d) (1, 10); e) (1, 10, 3).
5. Demostrar que un grupo con 4 o menos elementos es forzosamente abeliano. (*Sugerencia:* ba es uno de los e, b, a, ab, excepto casos triviales.)
6. Demostrar que si $xx=x$ en un grupo, será $x=e$.
7. ¿Cuáles de las siguientes tablas de multiplicación definen grupos?

	a	b	c	d
a	b	d	a	c
b	d	c	b	a
c	a	b	c	d
d	c	a	d	b

	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	a
d	d	c	b	b

8. Escribir la tabla de multiplicación para el grupo de las simetrías del triángulo equilátero.
9. ¿Cuáles de los siguientes conjuntos de números son grupos? a) Todos los números racionales, para la adición; para la multiplicación; b) Todos los números irracionales para la multiplicación; c) Todos los números complejos de valor absoluto 1, para la multiplicación; d) Todos los números complejos z con $|z|=1$, para la operación $z \cdot z'=|z| \cdot z'$; e) Todos los enteros, para la sustracción; f) Las «unidades» (Cap. 4, § 5) de cualquier dominio de integridad, para la multiplicación.
10. Demostrar que los siguientes postulados caracterizan un grupo abeliano: I) $(ab)c=a(cb)$ para a, b, c cualesquiera; II) El postulado de identidad a la izquierda del Teorema 4; III) El postulado de inverso a la izquierda del Teorema 4.

*11. Demostrar que si $x^2=e$ para todos los elementos de un grupo G, entonces G es conmutativo.

*12. Demostrar el Teorema 5. (*Sugerencia:* Si $ax=a$, x será elemento idéntico a la derecha, y cualquier identidad a la derecha será igualmente identidad a la izquierda.)

*13. Sea S un conjunto con una multiplicación tal, que $ab=ba$, $a(bc)=(ab)c$ y $ax=ay$ implica $x=y$: a) Si S es finito, demostrar que S es un grupo; b) Si S es finito o infinito, demostrar que S puede sumergirse en un grupo

*14. Demostrar que para la correspondencia $n \to n+1$ sobre la clase de los enteros positivos, hay inversos a la derecha, pero no inversos a la izquierda.

*15. Demostrar que cualquier inverso a la izquierda de una transformación de un conjunto finito sobre sí mismo, es también inverso a la derecha.

5. Isomorfismo

Consideremos la transformación $x \to \log x$ sobre el dominio de los números reales. Es bien sabido que cuando x crece en el intervalo $0 < x < +\infty$, $\log x$ crece continuamente en el intervalo $-\infty < y < +\infty$; o sea, que la correspondencia es biunívoca entre el sistema de los números reales positivos y el sistema de todos los números reales (la transformación inversa es $y \to e^y$). Además, $\log(xy) = \log x + \log y$, para valores cualesquiera de x e y; podemos reemplazar el cálculo de productos por el cálculo de las correspondientes sumas. ¡Ésta es, precisamente, la razón práctica del empleo de los logaritmos!

Sea J_3 el cuerpo de los enteros módulo 3 (Cap. I, § 10), y sea G el grupo de las rotaciones de un triángulo equilátero sobre sí mismo. Si I, R, R' son las rotaciones de 0°, 120°, 240°, respectivamente, la correspondencia $0 \leftrightarrow I$, $1 \leftrightarrow R$, $2 \leftrightarrow R'$, que asocia los enteros con las rotaciones, transporta las sumas de J_3 sobre los productos de las rotaciones correspondientes. Por ejemplo:

$$1+2 \equiv 0 \ (\text{mód. } 3), \qquad RR' = I$$

$$2+2 \equiv 1 \ (\text{mód. } 3), \qquad R'R' = R \qquad (240° + 240° \to 120°).$$

Lo anterior da ejemplos del concepto general de «isomorfismo», mencionado en el Cap. I, § 12. Este concepto es más simple y a la vez más importante para los grupos que para los dominios de integridad.

DEFINICIÓN. *Entenderemos por isomorfismo entre dos grupos* G *y* G′, *una correspondencia biunívoca* $a \leftrightarrow a'$ *entre sus elementos, que conserva la multiplicación del grupo; esto es, tal que* $a \leftrightarrow a'$ *y* $b \leftrightarrow b'$ *implica que* $ab \leftrightarrow a'b'$.

Así, en el primer ejemplo, hemos descrito un isomorfismo entre el grupo de los números reales positivos con la multiplicación, y el de todos los números reales con la adición. En el segundo hemos

señalado el isomorfismo del grupo aditivo de los enteros módulo 3 con el grupo de las simetrías por rotación del triángulo equilátero.

De la misma manera, el grupo de los enteros módulo 5 distintos de cero, con la multiplicación, es isomorfo con el grupo de enteros módulo 4, con la adición, siendo la correspondencia $1 \leftrightarrow 0$, $2 \leftrightarrow 1$, $4 \leftrightarrow 2$, $3 \leftrightarrow 3$.

Conviene comprobar este resultado comparando las tablas que dan las operaciones del grupo para los enteros módulo 4, con la adición, y para los enteros no nulos módulo 5, con la multiplicación. Estas tablas son :

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

×	1	2	4	3
1	1	2	4	3
2	2	4	3	1
4	4	3	1	2
3	3	1	2	4

Figura 4

A su vez, el grupo aditivo de los enteros mód. 4 es isomorfo con el grupo de rotaciones del cuadrado. Que la correspondencia $0 \leftrightarrow I$, $1 \leftrightarrow R$, $2 \leftrightarrow R'$, $3 \leftrightarrow R''$ es un isomorfismo, puede verse comparando la fig. 4 con parte de la fig. 3.

La noción de isomorfismo es muy importante porque formula la idea de que un mismo grupo abstracto puede aparecer en varias situaciones concretas de contextura diferente. El hecho de que grupos isomorfos sean abstractamente uno mismo (y que difieran sólo por la notación de sus elementos) puede verse de muchas maneras.

Así pues, por definición, dos grupos G y G' son isomorfos cuando cualquier tabla de multiplicar para G es tabla para G', con una sustitución apropiada. De aquí deduciremos la validez de lo dicho al final del § 4, o sea, que la condición necesaria y suficiente para que G' sea abeliano es que lo sea G, esto es, que cualquier imagen isomorfa de un grupo abeliano finito es un grupo abeliano.

Además, el isomorfismo viene a ser una especie de igualdad, en el aspecto que vamos a exponer :

TEOREMA 6. *La relación «G es isomorfo con G'» es una relación entre grupos, reflexiva, simétrica y transitiva.*

Demostración. La propiedad reflexiva es trivial (cada grupo es isomorfo consigo mismo por la transformación idéntica). Para la propiedad simétrica, sea $a \leftrightarrow aT$ una correspondencia isomorfa entre G y G'; como T es biunívoca, tiene una inversa T^{-1}, la cual es un isomorfismo que transporta a G' sobre G. Finalmente, si T transporta isomórficamente a G sobre G' y T' transporta isomórficamente a G' sobre G'', es claro que TT' será un isomorfismo entre G y G''.

Conviene observar que el Teorema 6 y su demostración son igualmente válidos para isomorfismos entre dominios de integridad, y también para isomorfismos entre sistemas algebraicos de cualquier clase que sean.

TEOREMA 7. *En cualquier isomorfismo entre dos grupos, los elementos idénticos se corresponden, y los elementos inversos de elementos correspondientes también son correspondientes.*

Demostración. La única solución e de $ax = a$ se corresponde con la única solución e' de $a'x = a'$; de aquí la correspondencia de los elementos idénticos. Consecuentemente, la única solución a^{-1} de la ecuación $ax = e$ en G, se corresponde con la única solución a'^{-1} de $a'x = e'$ en G'; esto completa la demostración.

Vamos finalmente a demostrar un importante resultado de Cayley, que puede interpretarse como una demostración de que el sistema de los precedentes postulados sobre la multiplicación de transformaciones es completo.

TEOREMA 8. *Todo grupo abstracto G es isomorfo con un grupo de transformaciones.*

Asociemos a cada elemento $a \in G$ la transformación ϕ_a : $x \rightarrow xa = x\phi_a$ sobre el «espacio» de todos los elementos x de G. Como $e\phi_a = e\phi_b$ implica $a = ea = eb = b$, elementos distintos de G se corresponden con distintas transformaciones. Como a la vez

$$(6) \qquad x(\phi_a\phi_b) = (x\phi_a)\phi_b = (xa)b = x(ab) = x\phi_{ab}$$

vale para todo x, el producto $\phi_a\phi_b$ es ϕ_{ab}, y el conjunto G' de todas las ϕ_a contiene con cada par de transformaciones su producto. Además, $x\phi_e = xe = x$ para todo x, luego G' contiene a la identidad. Puede verse de modo parecido que $(\phi_a)^{-1}$ existe y está en G', sien-

do precisamente $\phi(x^{-1})$. Así que, en resumen, G' es un grupo de transformaciones el cual, por (6), es isomorfo con G.

EJERCICIOS

1. ¿Son isomorfos dos cualesquiera de los siguientes grupos: a) el grupo de las simetrías de un triángulo equilátero; b) el grupo de las simetrías de un cuadrado; c) el grupo de las rotaciones de un hexágono regular; d) el grupo aditivo de enteros, mód. 6?
2. Resolver las mismas cuestiones para: a) Grupo de las rotaciones de un cuadrado; b) Grupo de las simetrías de un rectángulo; c) Grupo de las simetrías de un rombo; d) Grupo multiplicativo de 1, 5, 8, 12, mód. 13; e) Grupo multiplicativo de 1, 5, 7, 11, mód. 12.
3. a) Demostrar que el grupo aditivo de los enteros de Gauss, $m+n\sqrt{-1}$ $(m, n \in J)$ es isomorfo con el grupo multiplicativo de las fracciones racionales de la forma $2^n 3^m$ $(m, n \in J)$.
   b) Demostrar que ambos son isomorfos con el grupo de todas las traslaciones de un retículo rectangular indefinido.
*4. ¿Son isomorfos el grupo multiplicativo de los números reales no nulos y el aditivo de todos los números reales?
5. Determinar todos los isomorfismos entre el grupo aditivo de J_4 y el grupo de rotaciones del cuadrado.
6. a) Mostrar un isomorfismo entre el grupo del cuadrado y un grupo de transformación sobre los cuatro vértices, 1, 2, 3, 4, del cuadrado.
   b) Mostrar explícitamente en este isomorfismo la correspondencia de los inversos, según el Teorema 7.
7. Hacer lo mismo para el grupo de todas las rotaciones de un hexágono.
8. Ilustrar el Teorema 8 mostrando un grupo de transformaciones isomorfo con cada uno de los grupos siguientes:
   a) el grupo aditivo de todos los números reales;
   b) el grupo multiplicativo de todos los números reales no nulos;
   c) el grupo aditivo de los enteros, mód. 4.

6. Grupos cíclicos

En todo grupo, las *potencias* enteras a^m de cualquier elemento a perteneciente al grupo, pueden definirse separadamente para exponentes positivos, cero y negativos.

Si $m > 0$, definimos:

$$a^m = a \cdot a \ldots a \ (m \text{ factores}), \quad a^0 = e, \quad a^{-m} = (a^{-1})^m. \tag{7}$$

Las dos principales leyes del cálculo con exponentes son válidas:

$$a^r a^s = a^{r+s} \qquad (a^r)^s = a^{rs}. \tag{8}$$

Pero, en cambio, generalmente es $(ab)^r \neq a^r b^r$.

Si los dos exponentes r y s son positivos, la primera fórmula (8) se deduce directamente de la definición (7) (cfr. Cap. I, § 5). En otro caso, cuando uno de los exponentes r o s sea cero, las (8) son inmediatas; cuando r y s son ambos negativos, el resultado (8) sale directamente de la última parte de la definición (7). Queda el caso en que un exponente es negativo y el otro positivo, o sea, $r=-m$ y $s=n$, con $m>0$ y $n>0$. Entonces,

$$a^{-m}a^{n}=(a^{-1})^{m}a^{n}=(a^{-1}\ldots a^{-1})\cdot(a\ldots a).$$

Por la ley asociativa podemos reducir sucesivamente las a con sus inversas a^{-1}. En el caso $n \geqslant m$ quedará a^{n-m}, mientras que si $n<m$ quedarán algunos factores a la izquierda, $(a^{-1})^{m-n}$ o $a^{-(m-n)}$. En ambos casos obtenemos $a^{-m}a^{n}=a^{n+(-m)}$, c. q. d.

La segunda parte de (8) puede establecerse con mayor sencillez. Si s es positivo, tendremos, por la primera parte de (8),

$$a^{r}a^{r}\ldots a^{r} \ (s \text{ factores})=a^{r+\cdots+r}=a^{rs}.$$

Si s es negativo, podemos obtener un desarrollo similar sin más que observar que $(a^{r})^{-1}=a^{-r}$, sea r positivo, cero o negativo. Si s es cero, el resultado es inmediato.

DEFINICIÓN. *Se llama orden de un elemento* a *en un grupo* G *al menor entero positivo* m (*) *tal que* $a^{m}=e$; *si ninguna potencia de* a *es igual a la identidad, diremos que* a *tiene orden infinito. El grupo* G *es cíclico si contiene un elemento* x *tal, que todo otro elemento del grupo es una potencia de él; este elemento se llama generador del grupo.*

Por ejemplo, el grupo de todas las rotaciones de un cuadrado sobre sí mismo está constituído por las cuatro potencias R, R^{2}, R^{3} y $R^{4}=I$, siendo R la rotación de 90° en sentido de las agujas de un reloj. Este grupo puede igualmente ser engendrado por R^{3}, que es una rotación de 90° en sentido contrario al de las agujas de un reloj, pues $R^{2}=(R^{3})^{2}$, $R=(R^{3})^{3}$ e $I=(R^{3})^{4}$, que con R^{3} constituyen el grupo.

TEOREMA 9. *Si un elemento* a *engendra el grupo cíclico* G, *éste queda determinado por el orden de* a, *salvo isomorfismos. Preci-*

(*) Su existencia está asegurada por el principio de buena ordenación.

sando más; si el orden de a es infinito, G es isomorfo con el grupo aditivo de los enteros; si el orden de a es un entero finito n, G es isomorfo con el grupo de los enteros módulo n.

Demostración. Primeramente, $a^r = a^s$ si, y sólo si, se verifica que

$$e = a^r(a^s)^{-1} = a^r a^{-s} = a^{r-s} \qquad \text{por (8).} \tag{9}$$

Además, si $r \neq s$, o es $r > s$ o $s > r$; luego, si el orden de a es infinito, no podrá ser $a^{r-s} = e$, para ningún par de tales valores r, s; así que no habrá dos potencias de a iguales con exponentes distintos. Por otra parte, según (8) $a^s a^t = a^{s+t}$; por consiguiente, la correspondencia $a^s \to s$ es un isomorfismo entre G y el grupo aditivo de los enteros, lo que prueba la primera parte de nuestro teorema.

Si el orden de a es finito, el conjunto de los enteros t que hacen $a^t = e$ contiene al cero y, por (8), contiene a la suma y diferencia de dos cualesquiera de sus componentes. De aquí que por el Teorema 6 del capítulo I, $a^t = e$ si, y sólo si, t es múltiplo del orden n de a, y así, por (9), $a^r = a^s$ si $n \mid (r - s)$ y sólo en este caso; de otro modo, $a^r = a^s$ equivale a $r \equiv s$ (mód. n). Finalmente, por (8), $a^r a^s = a^{r+s}$; por consecuencia, con la correspondencia $a^r \to r$, G resulta isomorfo con el grupo aditivo de los enteros módulo n, como queríamos demostrar.

COROLARIO. *El número de elementos de un grupo cíclico G es igual al orden de un elemento generador de G; dos grupos cíclicos del mismo orden son isomorfos.*

EJERCICIOS

1. Utilizando las definiciones $a^1 = a$, $a^{m+1} = a^m a$, demostrar las leyes (8), para exponentes positivos, por inducción.
2. Demostrar que si $(ab)^n = a^n b^n$ para todo n, entonces G es conmutativo, y viceversa.
3. ¿Cuántos generadores diferentes tiene un grupo cíclico de orden 6?
*4. Demostrar que cualquier grupo conmutativo de orden 6 es cíclico.
5. ¿Es cíclico el grupo multiplicativo de 1, 2, ..., 6, mód. 7? ¿Y el de 1, 3, 5, 7, mód. 8? ¿Y el de 1, 2, 4, 5, 7, 8, mód. 9?
6. Si un grupo cíclico G está engendrado por el elemento a de orden m, demostrar que a^k engendra a G si, y sólo si, es m. c. d. $(k, m) = 1$.
7. Con las hipótesis del Ejerc. 6, hallar el orden de cualquier elemento a^k de G.
8. Hallar el orden de cualquier elemento del grupo del cuadrado.

9. a) Expresar todas las simetrías del cuadrado, como I, R, R^2, R^3, H, HR, HR^2, HR^3. (Esto significa que todo el grupo está «engendrado» por R y H.)
   b) Demostrar que toda la tabla del grupo puede obtenerse a partir de las tres «relaciones características» $R^4=I$, $H^2=I$, $RH=HR^3$. Dar ejemplos.
10. Demostrar cuántas «relaciones características» análogas a las del Ejercicio 9 pueden obtenerse para el grupo de simetrías de cualquier polígono regular. (Éste es el llamado «grupo del diedro».)
*11. Obtener los generadores y las relaciones características del grupo de simetrías de los tres modelos […] imaginando que se extienden indefinidamente en ambos sentidos. ¿Son isomorfos dos cualesquiera de estos tres grupos?
*12. Hacer un estudio similar para los grupos descritos en los ejercicios 1, 2, 4, 5, de § 3.

7. Grupos de sustituciones

Una *sustitución* es una transformación biunívoca de un conjunto finito en sí mismo. Se le llama también *permutación*.

Por ejemplo, el conjunto puede estar constituído por los cinco dígitos 1, 2, 3, 4, 5, y una sustitución puede ser la transformación ϕ,

$$(10)\qquad 1\phi=2,\quad 2\phi=3,\quad 3\phi=4,\quad 4\phi=5,\quad 5\phi=1.$$

Otra, puede ser la transformación ϕ' con

$$(11)\qquad 1\phi'=2,\quad 2\phi'=3,\quad 3\phi'=1,\quad 4\phi'=5,\quad 5\phi'=4.$$

El lector puede calcular $\phi\phi'$, $\phi'\phi$, y observará que $\phi\phi' \neq \phi'\phi$ (*).

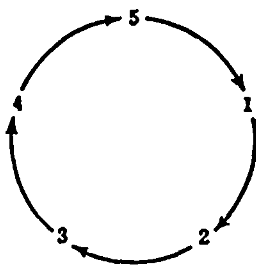


Figura 5

Las sustituciones que, como la ϕ definida antes, conservan una ordenación circular de los símbolos sustituídos, se llaman *ciclos* o *sustituciones cíclicas*. Para designarlas con una notación sencilla, se escribirán entre paréntesis primero una letra, detrás la que la sustituye, después la que sustituye a ésta, ..., y por último la que se transforma en la pri-

---

(*) Con una notación cómoda, y bastante frecuente, las sustituciones (10) y (11) se expresan, respectivamente,

$$\phi=\begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix},\qquad \phi'=\begin{pmatrix}1&2&3&4&5\\2&3&1&5&4\end{pmatrix}$$

N. del T.

mera. Así, la sustitución φ de (10) puede escribirse de una cualquiera de las maneras (12345), (23451), (34512), (45123) o (51234).

Teorema 10. *Una sustitución cíclica de n elementos tiene orden n.*

Demostración. La sustitución cíclica $\gamma = (a_1a_2 \ldots a_n)$ transporta a_i sobre a_{i+1}. Luego γ^2 tiene el efecto doble, y lleva cada a_i a a_{i+2} y generalmente, γ^k transporta a_i sobre a_{i+k} (en donde todos los subíndices deben reducirse módulo n). Tendremos en γ^k la identidad I cuando $a_{i+k} = a_i$; esto es, si $k \equiv 0$ (mód. n). El menor valor de k para $\gamma^k = I$ es entonces el mismo n, así que γ tiene orden n (ver definición en § 6). Se dice a veces que el ciclo es de *longitud n*.

La notación para una sustitución circular puede extenderse a cualquier sustitución. Por ejemplo, la sustitución de (11) permuta circularmente a los dígitos 1, 2, 3, entre sí, y asimismo 4 y 5. Es, pues, el producto de los dos ciclos (123)(45) = (45)(123). Este producto puede escribirse en cualquier orden, ya que los símbolos sustituídos en (123) quedan inalterados por (45), lo cual implica que la aplicación sucesiva de estas sustituciones, en cualquier orden, da el mismo resultado.

Teorema 11. *Toda sustitución φ puede escribirse como un producto de ciclos sin elementos comunes (brevemente: ciclos disjuntos).*

Demostración. Elijamos un elemento y llamémosle a_1. Designemos al $a_1\varphi$ por a_2, al $a_2\varphi$ por a_3, ..., al $a_{n-1}\varphi$ por a_n, hasta que $a_n\varphi = a_i$ coincida con alguno de los elementos ya nombrados. Como el precedente de cada a_i $(i > 1)$ es a_{i-1}, $a_n\varphi$ puede ser únicamente a_1. Entonces, el efecto de φ sobre las letras a_1, ..., a_n es el ciclo $(a_1a_2 \ldots a_n)$. Por otra parte, $(a_1 \ldots a_n)$ contiene, con cada símbolo a_i, su precedente; de aquí que φ permuta entre sí a los restantes elementos. El resultado se obtiene ya por inducción sobre el número de elementos. En particular, la sustitución idéntica sobre m letras se representa por m «ciclos» de longitud 1.

Recíprocamente, todo producto de ciclos sin elementos comunes representa una sustitución. Ahora podemos deducir que

Teorema 12. *El orden de cualquier sustitución es el mínimo común múltiplo de las longitudes de sus ciclos sin elementos comunes (disjuntos).*

Demostración. Escribamos la sustitución ϕ como el producto $\phi=\gamma_1\ldots\gamma_r$ de ciclos disjuntos γ_i. Como γ_i y γ_j son disjuntos, $\gamma_i\gamma_j=\gamma_j\gamma_i$ y los factores γ_i pueden ser permutados en ϕ y en sus potencias, para que sea $\phi^n=\gamma_1{}^n\ldots\gamma_r{}^n$ para todo n. Entonces, $\phi^n=I$ cuando cada $\gamma_r{}^n$ es la identidad. Pero por el Teorema 10 esto significa que $\phi^n=I$ cuando n sea múltiplo común de las longitudes de las γ_r, de lo cual se obtiene la conclusión del teorema.

Todo grupo finito es isomorfo con uno o más grupos de sustituciones, por el Teorema 8 del § 5. Por lo tanto, toda la teoría de grupos abstractos puede desarrollarse indirectamente, como una rama de la teoría de grupos de sustituciones. Aunque no lo haremos así, vale la pena presentar algunos ejemplos de grupos de sustituciones.

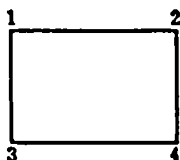


Figura 6

Consideremos el grupo de simetrías del rectángulo (fig. 6). Respecto a éstas, los vértices se transforman según las cuatro sustituciones

$$I=(1)(2)(3)(4), \qquad R=(14)(23), \qquad H=(13)(24), \qquad V=(12)(34).$$

A este grupo le llamaremos *grupo del rectángulo*. Conforme al Teorema 8, es isomorfo con el grupo de sustituciones: $\phi_I=(I)(R)(V)(H)$, $\phi_R=(IR)(HV)$, $\phi_H=(IH)(RV)$, $\phi_V=(IV)(RH)$.

El grupo de las simetrías del cuadrado (§ 1) puede representarse de un modo parecido como un grupo de sustituciones entre los cuatro vértices. Utilizando el Teorema 8, podemos también representarlo como un grupo de sustituciones entre los ocho símbolos que representan los elementos del grupo. Así, R corresponde a la sustitución efectuada en estos símbolos al multiplicarlos a la derecha por «R»; viendo la columna encabezada por «R» en la tabla de la figura 3, se observa que esta sustitución es $(IRR'R'')(HD'VD)$. Análogamente, H corresponde a $(IH)(RD)(R'V)(R''D)$.

El grupo (Teorema 2) de todas las sustituciones con n símbolos se llama *grupo simétrico* de grado n.

TEOREMA 13. *El grupo simétrico de grado* n *contiene* n! *sustituciones.*

En la construcción de una sustitución, la imagen k_1 del primer símbolo puede elegirse de n maneras, la del segundo símbolo puede

tomarse entre los $n-1$ símbolos distintos del k_1, y así sucesivamente. En total, resultan así $n(n-1)(n-2)\ldots 2\cdot 1 = n!$ sustituciones posibles.

EJERCICIOS

1. Expresar como un producto de ciclos disjuntos las permutaciones

   a) $1\varphi=4$, $2\varphi=6$, $3\varphi=5$, $4\varphi=1$, $5\varphi=3$, $6\varphi=2$;
   b) $1\varphi=5$, $2\varphi=3$, $3\varphi=2$, $4\varphi=6$, $5\varphi=4$, $6\varphi=1$;
   c) $1\varphi=3$, $2\varphi=5$, $3\varphi=6$, $4\varphi=4$, $5\varphi=1$, $6\varphi=2$.

   Hallar el orden de cada una de estas permutaciones.
2. Representar los productos siguientes como productos de ciclos disjuntos:

   $$(1234)(567)(261)(47); \quad (12345)(67)(1357)(163); \quad (14)(123)(45)(14).$$

   Hallar el orden de cada producto.
3. Hallar el orden de $(abcdef)(ghij)(klm)$; lo mismo con $(abcdef)(abcd)(abc)$.
4. Representar el grupo de las simetrías del rombo (paralelogramo equilátero) como un grupo de permutaciones entre sus vértices.
5. Lo mismo para el hexágono regular.
6. ¿Cuándo los grupos de simetrías son abelianos?
7. Sea G el grupo de todas las simetrías del cubo que dejan un vértice fijo. Representar G como un grupo de sustituciones entre los vértices (cfr. § 3).
8. a) Demostrar que cualquier permutación puede escribirse como un producto de «transposiciones» (esto es, ciclos de orden 2), en general no disjuntos.

   b) ¿Cómo se relaciona esto con la demostración de la «ley conmutativa generalizada» a partir de la ley $ab=ba$ (Cap. I, § 2)?
9. Representar el grupo de las simetrías de un triángulo equilátero como un grupo de permutaciones entre a) tres, y b) seis letras. En el caso b) síganse dos métodos esencialmente distintos, si se puede.
10. Demostrar que el grupo simétrico de grado n es engendrado por los ciclos $(1, 2, \ldots, n-1)$ y $(n-1, n)$.

*11. ¿En qué sentido es única la representación del Teorema 11?

8. Subgrupos

Muchos grupos están contenidos en otros más amplios, así el grupo de las rotaciones del cuadrado es una parte del grupo de las simetrías del cuadrado. Además, el grupo de las ocho sustituciones entre los vértices del cuadrado que determinan las simetrías, es una parte del grupo de las $4!=24$ sustituciones de estos vértices. El grupo de los números enteros pares, con la adición, es una parte del grupo que constituyen todos los enteros, con la adición.

Estos ejemplos sugieren el concepto de subgrupo. Un subconjunto S de un grupo G se llama un *subgrupo* de G si también S es un grupo, con respecto a la misma operación binaria (multiplicación) de G.

En todo grupo G, el conjunto formado por sólo la identidad e es un subgrupo. La totalidad de G es también un subgrupo de G. Pero estos dos subgrupos son triviales (llamados «impropios»). Se llama subgrupos propios de G a los restantes, esto es, los distintos de e y G.

TEOREMA 14. *Para que un subconjunto* S *no vacío de un grupo* G *sea un subgrupo de éste, es necesario y suficiente que 1) si* a *y* b *están en* S, *lo esté* ab : *2) si* a *está en* S, *también lo esté* a^{-1}.

Con estas hipótesis, S es claramente un subgrupo; la asociatividad es trivial; la identidad $e=aa^{-1}$ de G pertenece a S, por haber en S al menos un elemento a; los restantes postulados del grupo están implicados en la hipótesis. Recíprocamente, debemos probar que 1) y 2) se verifican en cualquier subgrupo. La identidad $x=e'$ de todo subgrupo de G satisface a $xx=x$, luego es también la identidad de G (Ejerc. 6, §4). Por consecuencia, como en G hay un solo inverso de cada a, el inverso de cada elemento a en el subgrupo es el mismo inverso que en G, y se verifica 2). La condición 1) es obvia.

Para elementos a de orden finito m, $a^{m-1}a=a^m=e$ y entonces $a^{-1}=a^{m-1}$. De aquí se obtiene la siguiente condición, más sencilla:

TEOREMA 15. *Un subconjunto* S *no vacío de un grupo finito* G *es un subgrupo de* G *cuando el producto de dos elementos cualesquiera de* S *está también en* S.

El problema de la determinación de todos los subgrupos de un grupo dado G es en general muy difícil. Veremos ahora cómo se resuelve en el caso de que G sea un grupo cíclico.

TEOREMA 16. *Cualquier subgrupo* S *de un grupo cíclico* G *es también cíclico.*

Demostración. Sabemos que G está constituido por las potencias de uno de sus elementos a. Si a^s y a^t están en S, también lo estarán $a^{s+t}=a^sa^t$ y $a^{s-t}=a^s(a^t)^{-1}$, por el Teorema 14. El conjunto

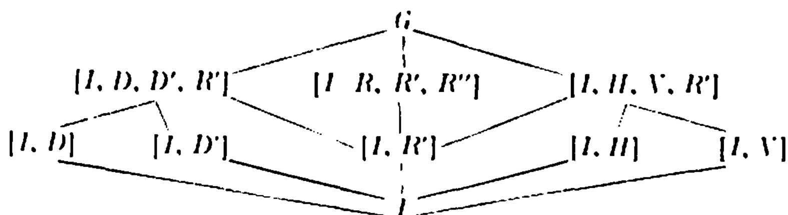
de los enteros s para los que a^s está en S es, pues, un conjunto cerrado para la adición y la sustracción, así que estará compuesto por los múltiplos de un exponente mínimo r (Teorema 6, Cap. I). Por lo tanto, S consistirá en el conjunto de las potencias $a^{kr} = (a^r)^k$, luego es cíclico, y a^r es su elemento generador (*).

En el caso de ser G infinito, cada $r > 0$ determina un subgrupo diferente. Si G tiene n elementos, como $a^n = e$ está seguramente en S, únicamente determinan subgrupos aquellos valores $r > 0$ que son divisores de n, y además, estos subgrupos son distintos.

Para disponer de ejemplos, con vista a ulteriores desarrollos, vamos a obtener todos los subgrupos del grupo del cuadrado. Recordando las definiciones dadas en § 1 para las operaciones de este grupo, se encuentran los subgrupos propios, que dejan respetivamente invariante cada una de las siguientes configuraciones:

una diagonal	*un eje*	*una cara*	*un eje y una diagonal*
$[I, D, D', R']$	$[I, H, V, R']$	$[I, R, R', R'']$	$[I, R']$
vértice 1 (ó 3)	*vértice 2 (ó 4)*	*lados verticales*	*lados horizontales*
$[I, D]$	$[I, D']$	$[I, H]$	$[I, V]$

Por transformaciones que dejan invariante una cara, entendemos aquellas que no hacen salir al cuadrado de su plano. Todos estos subgrupos pueden ser representados en sus mutuas relaciones en un esquema donde cada grupo está unido por líneas descendentes con todos sus subgrupos.



Sin recurrir a consideraciones geométricas podemos también hallar todos estos subgrupos. La determinación de todos los subgrupos de un grupo finito es más manejable considerando el grupo de elementos abstractos.

(*) Este razonamiento sólo podría dar $r = 0$ si el conjunto S consistiese en I solo.

El procedimiento usual es observar primero que si un subgrupo S de G contiene un elemento a ha de contener también el subgrupo cíclico $\{a\}$ (pruébese que es un subgrupo) formado por todas las potencias de a. En el caso anterior, esta consideración proporciona todos los subgrupos, excepto los dos extremos de la primera fila. Notemos después que todo subgrupo que contenga dos subgrupos cíclicos $\{a\}$ y $\{b\}$ ha de contener el conjunto $\{a, b\}$ de todos los productos, tales como $a^2b^{-3}a$, formados por potencias de a y b. (Demostrar, haciendo uso del Teorema 15, que estos productos forman un subgrupo.) En el caso antedicho, se obtienen así los restantes subgrupos (veremos en § 9 por qué todos los subgrupos contienen 2 ó 4 elementos). En general, deberemos considerar después los subgrupos $\{a, b, c\}$ engendrados por tres o más elementos; pero esto solamente es necesario hacerlo cuando el número de elementos del grupo sea un producto de cuatro o más factores primos.

La *intersección* $S \cap T$ de dos subgrupos S y T (¡y también de dos conjuntos cualesquiera!) es el conjunto de todos los elementos que pertenecen a la vez a S y a T.

Teorema 17. *La intersección* $S \cap T$ *de dos subgrupos* S *y* T *de un grupo* G*, es un subgrupo de* G.

Por el teorema 14, a en $S \cap T$ implica a en S, y de aquí a^{-1} en S; del mismo modo supone a^{-1} en T, luego a^{-1} en $S \cap T$. Análogamente, a y b en $S \cap T$ implica ab en S y ab en T, y por tanto, ab en $S \cap T$. Por ello, y según el Teorema 14, $S \cap T$ es un subgrupo. Además, $S \cap T$ contiene a e, y por tanto no es vacío.

Claramente se ve que $S \cap T$ es el subgrupo más extenso contenido a la vez en S y en T; dualmente, existe un subgrupo mínimo que contiene a la vez a S y a T. Está constituído por los productos de potencias positivas y negativas de elementos de S y T y se le llama la unión de S y T y se le representa por $S \cup T$. Recurriremos a estos conceptos en el cap. XI.

EJERCICIOS

1. En el grupo de simetrías del hexágono regular ¿qué subgrupos dejan una diagonal fija?
2. Si T es un subgrupo de S, y S un subgrupo de G, demostrar que T es un subgrupo de G.
3. En el grupo de todas las permutaciones ϕ entre cuatro elementos 1, 2, 3, 4, hallar los siguientes subgrupos: a) todas las ϕ que transforman el

conjunto (1, 2) en el mismo conjunto (1, 2); b) todas las ϕ tales, que $a \equiv b$ (mód. 2) implica $a\phi \equiv b\phi$ (mód. 2) para todos los números a y b del conjunto 1, 2, 3, 4. (*Sugerencia:* (13)(24) es una permutación ϕ.)

4. Demostrar que el Teorema 15 vale también si G es infinito, pero todos los elementos de G tienen orden finito. Dar el ejemplo de un grupo de esta especie.
5. Tabular todos los subgrupos de los grupos siguientes: a) el grupo aditivo, mód. 12; b) el grupo de un pentágono regular; c) el grupo de un hexágono regular; *d) el grupo de todas las permutaciones entre cuatro letras.
* 6. Sea $a \leftrightarrow a'$ un isomorfismo entre dos grupos G y G' de permutación, y sea S el conjunto de aquellas permutaciones de G que tienen una letra invariante. ¿Debe ser necesariamente un subgrupo de G' el conjunto S' de aquellos elementos de G', correspondientes a los a de S?
7. Demostrar que, en cualquier grupo G, el conjunto de los elementos a tales que $ax = xa$ para todo $x \in G$, es un subgrupo de G (éste es el llamado «centro» de G).
8. Hallar el centro (ejerc. 7) del grupo del cuadrado, y el del grupo simétrico con tres letras.
* 9. Hacer lo mismo, a) para el grupo de un polígono regular de n lados; b) para el grupo simétrico con n letras.

9. Cogrupos o clases de restos. Teorema de Lagrange

Vamos a establecer ahora un concepto de gran trascendencia en la teoría de grupos abstractos: la idea de que cualquier subgrupo S de un grupo G descompone a G en cogrupos (*) (o en clases de restos, como también se dice).

DEFINICIÓN. *Llamaremos orden de un grupo o de un subgrupo al número de sus elementos. Llamaremos cogrupo a la derecha (cogrupo a la izquierda) de un subgrupo* S *en un grupo* G, *al conjunto* Sa *(respectivamente,* aS) *de todos los múltiplos a la derecha* Sa *(a la izquierda,* aS) *de los elementos* s *de* S, *por un elemento fijo* a *de* G. *El número de cogrupos a la derecha de* S *se llama el índice de* S *en* G.

Como $Se = S$, S es por sí mismo un cogrupo a la derecha. Además se tiene:

LEMA 1. *Si* S *es finito, cada cogrupo a la derecha de* S *tiene exactamente el mismo número de elementos que* S.

(*) La locución *cogrupo* no debe inducir al error de suponer que sus elementos constituyan un subgrupo: esto no sucede nunca, excepto para $Se = S$. (N. del T.)

Por ser la transformación $s \to sa$ biunívoca, cada elemento $t = sa$ del cogrupo Sa es la imagen de un elemento $s = ta^{-1}$ de S, y sólo de uno. (Cfr. también el Teorema 8.)

LEMA 2. *Dos cogrupos a la derecha* Sa *y* Sb *de* S, *o son idénticos o no tienen ningún elemento común.*

En efecto: supongamos que Sa y Sb tengan un elemento común $c = s'a = s''b$ (s' y s'' en S). Entonces Sb contiene a todos los elementos $sa = ss'^{-1}s'a = (ss'^{-1}s'')b$ de Sa, y, análogamente, Sa contiene a todos los elementos de Sb. Por consiguiente, $Sa = Sb$.

Es fácil ilustrar con ejemplos estos resultados. Así, si G es el grupo del cuadrado, el subgrupo $S = [I, H]$ tiene cuatro cogrupos a la derecha:

$$[I, H]I = [I, H]; \qquad [I, H]R = [R, HR] = [R, D'];$$
$$[I, H]R' = [R', HR'] = [R', V]; \qquad [I, H]R'' = [R'', HR''] = [R'', D].$$

Cada cogrupo tiene dos elementos y cualquier elemento del grupo está en uno de estos cuatro cogrupos.

Si G es el grupo aditivo de los enteros, el subgrupo constituído por los múltiplos de 5 tiene por cogrupos a la derecha las diferentes clases residuales módulo 5. Finalmente, sea G el grupo simétrico de todas las sustituciones entre los símbolos 1, ..., 6, y sea S el subgrupo constituído por los elementos de G que dejan fijo el símbolo 1. Entonces $1\phi = K$ implica para todo $\psi \in S$ que $1(\psi\phi) = (1\psi)\phi = 1\phi = K$. De aquí que el cogrupo S contiene solamente las permutaciones que transportan 1 sobre K (y, por el Lema 1, las contiene a todas). Así pues, los cogrupos a la derecha de S son los subconjuntos que transportan $1 \to 1$, $1 \to 2$, ..., $1 \to 6$, respectivamente. Cada cogrupo consta de 5! elementos.

Como cada cogrupo Sa contiene siempre al elemento $a = ea$, todo el grupo G se agota en sus cogrupos a la derecha. En cuyo caso, G queda descompuesto en subconjuntos disjuntos, cada uno de los cuales tiene tantos elementos como S. Si G es finito (*), la conclusión es:

TEOREMA 18 (Lagrange). *El orden de un grupo finito* G *es múltiplo del orden de cualquiera de sus subgrupos.*

(*) La extensión al caso infinito es posible (Cap. XII, § 4, Ejerc. 10), pero no importante.

Cada elemento a de G engendra un subgrupo cíclico cuyo orden es (Teorema 9) el orden de a. Así tendremos el

COROLARIO 1. *El orden de cualquier elemento de un grupo finito G es divisor del orden de G.*

COROLARIO 2. *Todo grupo G de orden primo p es cíclico.*

Porque el subgrupo cíclico $\{a\}$ engendrado por un elemento $a \neq e$ en tal grupo, tiene un orden $n > 1$ divisor de p. Esto implica $n = p$, así que $G = \{a\}$ es cíclico.

COROLARIO 3. *Los únicos grupos abstractos de orden cuatro son el grupo cíclico de orden cuatro y el grupo del rectángulo.*

Si un grupo G de orden 4 contiene un elemento de orden 4, es cíclico. Por otra parte, por el Corolario 1, todos los elementos de G excepto e han de tener orden 2. Llamémoslos a, b, c. Por la ley de simplificación, ab no puede ser igual a $ae = a$, $eb = b$ ni $aa = e$, luego $ab = c$. Por simetría, $ac = ca = b$, $bc = cb = a$, $ba = c$. Pero esto, unido a $a^2 = b^2 = c^2 = e$ y $ex = xe = x$, nos da la tabla de multiplicar del grupo del rectángulo, como se vió en § 7.

El teorema de Lagrange puede aplicarse también a la teoría de los números.

COROLARIO 4 (Fermat). *Si* a *es entero y* p *es primo, resulta* $a^p \equiv a$ (*mód.* p).

El grupo multiplicativo módulo p (excluyendo al cero) tiene $p-1$ elementos. El orden de cualquier elemento a de estos grupos es un divisor de $p-1$, por el Corolario 1; así que $a^{p-1} \equiv 1$ (mód. p), cualquiera que sea $a \not\equiv 0$ (mód. p). Si se multiplican por a los dos miembros obtendremos la congruencia que deseamos, excepto para el caso $a \equiv 0$ (mód. p), en el cual la conclusión es una verdad trivial.

EJERCICIOS

1. Comprobar el teorema de Fermat para $p=7$ y $a=2, 3, 6$.
2. a) Enumerar los subgrupos del grupo del diedro (§ 6, Ejerc. 10) de orden 26. ¿Cuántos existen?
   b) Generalizar el resultado.
3. Demostrar: el número de cogrupos a la derecha de cualquier subgrupo de un grupo finito es igual al número de sus cogrupos a la izquierda. (*Sugerencia:* Utilizar la correspondencia $x \to x^{-1}$.)
4. Determinar los cogrupos del subgrupo [*I*, *D*] del grupo del cuadrado.

5. Si S es cualquier subgrupo de un grupo G, denotemos por SaS el conjunto de todos los productos asa' para s, s' en S. Demostrar que, o bien $SaS \frown SbS$ es vacío, o bien $SaS = SbS$, para todo $a, b \in G$.
6. Para un subgrupo S, definamos $x \equiv y$ (mód. S), para significar que $xy^{-1} \in S$.
   a) Demostrar que esta relación es reflexiva, simétrica y transitiva, y mostrar que $x \equiv y$ (mód. S) si, y sólo si, x e y pertenecen al mismo cogrupo a la derecha de S.
   b) Demostrar que $x \equiv y$ (mód. S) implica $xa \equiv ya$ (mód. S) para todo a.
7. Sea G el grupo de un hexágono regular, S el subgrupo que deja un vértice fijo. Hallar los cogrupos a la izquierda y a la derecha de S.
8. Describir los cogrupos por la derecha y por la izquierda del subgrupo de todas las permutaciones de $x_1, \dots, x_4$ que transporten sobre sí mismo el conjunto $[x_1, x_2]$.
9. Demostrar que un grupo de orden p^m, con p primo, debe contener un subgrupo de orden p.
10. a) Si G es el grupo de todas las transformaciones $x \to ax+b$ de R^*, con $a \neq 0$ y b real, mientras que S es el subgrupo de todas estas transformaciones con $a=1$, describir los cogrupos de S en G.
    b) Hacer lo mismo para el subgrupo T de todas las transformaciones con $b=0$.
11. a) Demostrar que para cualquier entero $n > 1$, los enteros positivos $k < n$, primos con n, forman un grupo G para la multiplicación mód. n.
    b) Demostrar que si $\phi(n)$ denota el orden de G, entonces $x^{\phi(n)} \equiv 1$ (mód. n) para todo $x \in G$.
    c) Calcular $\phi(n)$ para $n = 12, 16, 30$.
*12. Si S y T son subgrupos de orden s y t en un grupo G, y si u y v son los órdenes de $S \frown T$ y de $S \smile T$, demostrar $st \leqslant uv$.
*13. Demostrar que los únicos grupos abstractos de orden 6 son el grupo cíclico y el grupo simétrico de tres letras.
*14. Sea 2^h+1 igual a un primo p.
    a) Demostrar que en el grupo multiplicativo mód. p, el orden de 2 es $2h$.
    b) Mediante el Teorema de Fermat, deducir que $2h$ divide a $p-1=2^h$.
    c) Concluir que h es una potencia de 2.

10. Sustituciones pares e impares

Una clasificación importante de las sustituciones resulta considerando la forma polinómica homogénea $P = \prod_{i<j} (x_i - x_j)$, en la que i y j toman todos los valores de 1 a n. Si $n=3$, P es:

$$(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) =$$
$$= x_1^2 x_2 + x_2^2 x_3 + x_3^2 x_1 - x_1^2 x_3 - x_3^2 x_2 - x_2^2 x_1.$$

En general, P es un polinomio de grado $n(n-1)/2$. Evidentemente, toda sustitución aplicada a los subíndices de P deja al conjunto de factores de P, y a P mismo, invariable, excepto el signo.

Así, la transposición (x_1x_2) cambia (x_1-x_2) en su opuesto (x_2-x_1), intercambia (x_1-x_j) y (x_2-x_j) $(j>2)$, y deja inalterados a los otros factores. En resumen, cambia P en $-P$.

Las $n!$ sustituciones de los subíndices son de dos clases: las sustituciones *pares*, que dejan P (y también $-P$) invariante, y las sustituciones *impares*, que intercambian P y $-P$. De aquí deducimos, cuando consideramos el efecto de dos sustituciones aplicadas sucesivamente, que se tiene:

(12)
$$\text{Par}\times\text{Par}=\text{Impar}\times\text{Impar}=\text{Par},$$
$$\text{Par}\times\text{Impar}=\text{Impar}\times\text{Par}=\text{Impar}.$$

Como corolario de (12) y del Teorema 15 resulta que las sustituciones pares forman un subgrupo A_n del grupo simétrico de grado n. A este subgrupo se le llama el «grupo alternado» de grado n.

Si β es una sustitución impar fija y ϕ una sustitución impar variable, entoces $\phi\beta^{-1}$ es par y $\phi=(\phi\beta^{-1})\beta$ está en el cogrupo a la derecha $A_n\beta$. En resumen, las permutaciones impares forman un solo cogrupo a la derecha de A_n. De aquí, por el teorema de Lagrange, que el «grupo alternado» de n símbolos contiene $(n!)/2$ elementos.

Un polinomio $g(x_1, x_2, \ldots, x_n)$, con n indeterminadas, se llama «simétrico» si es invariante en todas las sustituciones entre sus subíndices, es decir, cuando es invariante para el grupo simétrico. Polinomios simétricos son, en particular (para $n=3$):

(13)
$$\sigma_1=x_1+x_2+x_3,\quad \sigma_2=x_1x_2+x_1x_3+x_2x_3,\quad \sigma_3=x_1x_2x_3.$$

Estos son los coeficientes del desarrollo

(14)
$$(t-x_1)(t-x_2)(t-x_3)=t^3-\sigma_1t^2+\sigma_2t-\sigma_3.$$

En general llamaremos a tales polinomios, *polinomios simétricos elementales* (en n variables); son:

(15)
$$\sigma_1=\sum_i x_i,\quad \sigma_2=\sum_{i<j} x_ix_j,\quad \sigma_3=\sum_{i<j<k} x_ix_jx_k\ldots,\quad \sigma_n=x_1\ldots x_n.$$

Como $(-1)^k\sigma_k$ es el coeficiente de t^{n-k} en el desarrollo de $p(x)=$ $=\Pi_k(t-x_k)$, las expresiones σ_i dan los coeficientes de $p(x)$ como funciones de sus raíces. Una gran parte de su importancia se deriva del llamado «teorema fundamental de los polinomios simétricos», que vamos a enunciar sin demostración (véase, no obstante, Capítulo XV, Teor. 10, Corolario).

TEOREMA 19. *Todo polinomio simétrico* $p(x_1, \ldots, x_n)$ *puede expresarse como un polinomio de los polinomios simétricos elementales.*

Así, en el caso de dos variables x e y,

$$x^2+y^2=(x+y)^2-2xy=\sigma_1^2-2\sigma_2,$$
$$x^3+y^3=(x+y)^3-3xy(x+y)=\sigma_1(\sigma_1^2-3\sigma_2),$$

y así sucesivamente. Aunque un polinomio $g(x_1, \ldots, x_n)$ no sea simétrico, podemos preguntarnos cuál es el conjunto de las sustituciones entre sus índices que lo dejan invariable. Este conjunto de sustituciones es evidentemente un grupo, que se llama el "grupo del polinomio."

EJERCICIOS

1. Enumerar todas las permutaciones impares, a) de tres letras; b) de cuatro letras.
2. ¿Para qué enteros positivos n existe un ciclo de longitud n par? ¿E impar?
3. Dar un criterio para averiguar mentalmente cuándo un producto de ciclos es par o impar. Aplicarlo a (123)(246)(5432) y a (12)(345)(67)(891).
4. a) Construir ejemplos de permutaciones pares e impares de orden 14 con once letras.
   b) Demostrar que cualquier permutación de orden 10 con ocho letras es impar.
5. Demostrar que cualquier permutación par puede escribirse como el producto de ciclos de tercer orden.
6. Demostrar que una permutación es par si, y sólo si, puede escribirse como producto de un número par de transposiciones (Ejerc. 8, § 7).
7. Hallar el grupo de cada uno de los siguientes polinomios:

$$x_1x_2+x_3x_4, \quad x_2x_1+x_3x_2+x_3x_1, \quad x_1^2x_2+x_3x_4^2+x_1^2x_3+x_2x_4^2.$$

8. Representar cada uno de los siguientes polinomios en función de los polinomios simétricos elementales:

$$x^2+y^2+z^2, \quad x^2y+y^2z+z^2x+x^2z+y^2x+z^2y.$$

11. Elementos conjugados. Automorfismos

Existe una evidente semejanza entre ciertos pares de simetrías del cuadrado; así es, en particular, entre las reflexiones H y V, entre las reflexiones D y D' y entre las dos rotaciones R y R', dando a estas letras el significado explicado en § 1.

Estas tres analogías son debidas a la misma causa. Para darnos cuenta de ella, observemos las siguientes igualdades (confróntense en el cuadro de § 4):

$$V = R^{-1}HR, \qquad D' = R^{-1}DR, \qquad R' = H^{-1}RH.$$

Así, una transformación de cada uno de tales pares es «transformada» de la otra, según la siguiente

DEFINICIÓN. *En un grupo* G, *se llama transformado de* x *mediante* b *al elemento* $b^{-1}xb$. *A los elementos* $b^{-1}xb$ *se les llama también conjugados del* x ($x \in G$, $b \in G$).

La denominación de «transformado de x mediante b» se justifica con la siguiente interpretación geométrica: sean A y X transformaciones biunívocas en un espacio S; entonces, $Y = A^{-1}XA$ es análogo a X, en el siguiente sentido geométrico de la frase. Para todos los puntos $q = (pA)$ en S tenemos:

$$(pA)Y = (pA)(A^{-1}XA) = (pAA^{-1})XA = (pX)A.$$

Así, Y es la transformación $(pA) \to (pX)A$. ¡$Y = A^{-1}XA$ es, en fin, lo que viene a ser X si cada punto de S se transforma por A! Por ejemplo, la relación $V = R^{-1}HR$ expresa el hecho de que R transforma el eje horizontal en vertical, y por lo tanto, la reflexión en un eje horizontal se transforma en la reflexión en un eje vertical. Las igualdades $D' = R^{-1}DR$, $R' = H^{-1}RH$ y otras análogas tienen una significación similar. Pero la relación de elemento transformado (o conjugado) de otro es importante también para los grupos abstractos.

DEFINICIÓN. *Un isomorfismo de un grupo* G *consigo mismo se llama automorfismo de* G. *Así pues, un automorfismo* a *de* G *es una transformación biunívoca de* G *tal, que*

(16) $$(xy)a = (xa)(ya) \quad \textit{para todo } x, y, \textit{ de } G.$$

TEOREMA 20. *Para cada elemento fijo* a *del grupo* G, *la correspondencia* $T_a : x \to a^{-1}xa$ *es un automorfismo de* G.

Demostración. $(a^{-1}xa)(a^{-1}ya) = a^{-1}(xy)a$ para todo x, y.

Los automorfismos T_a de la forma $x \to a^{-1}xa$ se llaman *automorfismos internos*; los demás son los automorfismos *externos*.

TEOREMA 21. *Los automorfismos de un grupo G constituyen a su vez un grupo A.*

Demostración (cfr. Teorema 6). Es obvio que la transformación idéntica es un automorfismo y que también lo es el producto de dos automorfismos. Finalmente, si $x \to xa$ es un automorfismo, se tendrá por (16):

$$(xy)a^{-1} = [(xa^{-1}a)(ya^{-1}a)]a^{-1} = ([(xa^{-1})(ya^{-1})]a)a^{-1} = (xa^{-1})(ya^{-1}),$$

así que a^{-1} es un automorfismo, lo que demuestra el teorema.

Una definición y teorema análogos se aplican a los dominios de integridad, y también a las álgebras abstractas en general. Se puede considerar útilmente que un automorfismo de un álgebra abstracta A es una especie de simetría de A.

Puede comprobarse que el grupo de simetrías del cuadrado tiene cuatro automorfismos internos diferentes y otros cuatro externos. Por otra parte, el grupo cíclico de tercer orden no tiene automorfismos internos excepto la identidad, pero tiene el automorfismo «externo» $x \to x^2$.

TEOREMA 22. *Los automorfismos internos de un grupo G constituyen un subgrupo del grupo de los automorfismos de G.*

Demostración. Como $b^{-1}(a^{-1}xa)b = (ab^{-1})x(ab)$, el producto de dos automorfismos internos T_a y T_b es un automorfismo interno T_{ab}; además, siendo $(a^{-1})^{-1}(a^{-1}xa)(a^{-1}) = x$, el inverso del automorfismo T_a es $T_{(a^{-1})}$.

DEFINICIÓN (Galois). *Un subgrupo S de un grupo G se llama normal (en G) cuando es invariante para todos los automorfismos internos de G (o sea, que contiene, con cada elemento, a todos sus conjugados).*

En vez de subgrupo normal, se dice a veces subgrupo «invariante».

Así, el grupo de las rotaciones del cuadrado es un subgrupo normal del de todas sus simetrías; también lo es el subgrupo $[I, R^2]$. Cualquier subgrupo de un grupo abeliano es normal, ya que $a^{-1}xa = a^{-1}ax = x$ para todo a, x. El grupo de las traslaciones del plano es también un subgrupo normal del de todos los movimientos rígidos del plano (Cap. IX).

En general, designaremos con $a^{-1}Sa$ el conjunto de todos los productos $a^{-1}sa$ para todo s en S. La definición anterior establece que S es normal cuando el conjunto $a^{-1}Sa$ es igual al S, para todo a en G.

TEOREMA 23. *Un subgrupo* S *es normal cuando todos sus cogrupos a la derecha son cogrupos a la izquierda, y recíprocamente.*

Si S es normal, $aSa^{-1}=(a^{-1})^{-1}Sa^{-1}=S$ para todo a; de aquí que el conjunto Sa de los sa $(s \in S)$ es el mismo que el conjunto $(aSa^{-1})a$ de $(asa^{-1})a=as$ $(s \in S)$. Así, $Sa=aS$ para todo a. Recíprocamente, si el cogrupo a la derecha Sa es un cogrupo a la izquierda bS, entonces $a^{-1}Sa=a^{-1}bS$ contiene $e=a^{-1}ea$ y resulta (lema 2, § 9) igual a $eS=S$, lo cual completa la demostración.

Vemos, como corolario, que todo subgrupo S que tenga un solo cogrupo es normal; los elementos que no están en S forman el cogrupo por la derecha y por la izquierda de S. Se deduce que el grupo alternado es un subgrupo normal del grupo simétrico de grado n.

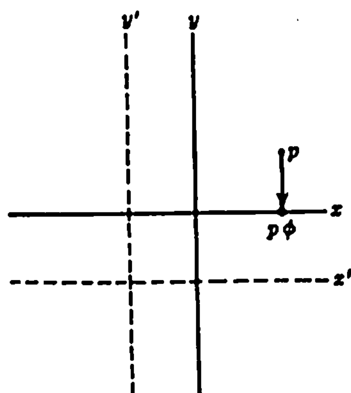


Figura 7

Observación. El concepto de «transformada» se aplica también a correspondencias sin inversa. Por ejemplo, la correspondencia $(x, y) \to (x, 0)$ representa una proyección ϕ del plano sobre el eje de las x (figura 7). Si reemplazamos, respectivamente, el eje de las x y el de las y por el $x' : y=-1$ y el $y' : x=-1$, obtendremos una transformación de coordenadas ψ. (Para este concepto básico, consúltese una Geometría Analítica.) Esta ψ asigna a cada punto nuevas coordenadas $(x', y')=(x+1, y+1)$; su inversa ψ^{-1} tendrá la forma $(x, y)=(x'-1, y'-1)$. Con relación a los nuevos ejes, la proyección ϕ tendrá la representación coordenada $\phi' : (x', y') \to (x', 1)$. Pero ésta es simplemente la $\psi^{-1}\phi\psi$, transformada de ϕ por ψ:

$$(x, y) \to (x-1, y-1) \to (x-1, 0)$$
$$\to (x-1+1, 0+1)=(x, 1).$$

EJERCICIOS

1. ¿Cuántos automorfismos tiene un grupo cíclico de orden p? ¿Y uno de orden pq? (p, q, primos distintos.)
2. Enumerar todos los automorfismos del grupo del rectángulo. ¿Cuáles son internos?
3. Calcular $t^{-1}at$ para la sustitución $a=(1234)$ y t igual, sucesivamente, a (13), (12)(34), (124), (1423). Establecer una regla general para calcular las sustituciones conjugadas de una dada.
4. Sea R la rotación $x'=x\cos\theta-y\operatorname{sen}\theta$; $y'=x\operatorname{sen}\theta+y\cos\theta$. Calcular la transformada de φ por R, si a) $(x, y)\phi=(x+1, y)$; b) $(x, y)\phi=(x+1, y+1)$.
5. Demostrar que en cualquier grupo la relación «x es conjugada con y» es una relación de equivalencia.
6. Demostrar que un elemento a de un grupo induce al automorfismo interno idéntico si, y sólo si, pertenece al «centro» (§ 8, Ejerc. 7).
7. Demostrar que cualquier subgrupo de un grupo cíclico es normal. Generalizar este resultado.
8. Demostrar que si G y H son grupos isomorfos, el número de isomorfismos distintos entre G y H es el número de automorfismos de G.
9. Enumerar los automorfismos internos, conjuntos de elementos conjugados, y subgrupos normales, del grupo del cuadrado.
10. Hacer lo mismo para el grupo simétrico con cuatro letras.
11. Hacer lo mismo para el grupo del hexágono regular.
12. Demostrar que si M y N son subgrupos normales de un grupo G, también lo será su intersección.
13. Demostrar que bajo las hipótesis del Ejerc. 12, el conjunto MN de todos los productos xy ($x\in M$, $y\in N$) es un subgrupo normal de G.
14. Demostrar que los automorfismos internos de cualquier grupo G son un subgrupo normal del grupo de todos los automorfismos de G.

*15. a) Demostrar que para cualquier $c\neq 0$, racional, la correspondencia $x\rightarrow xc$ es un automorfismo del grupo aditivo de todos los números racionales.
  b) Mostrar que este grupo no tiene otros automorfismos.

16. Sea G un grupo de orden pq (p, q primos). Demostrar que G, o bien es cíclico, o bien contiene un elemento de orden p (o q). En el segundo caso, demostrar que G contiene, o bien un subgrupo normal, o bien q subgrupos conjugados de orden p. En el último caso, mostrar los $pq-q(p-1)=q$ elementos de orden distinto de p que forman un subgrupo normal. Deducir que G tiene siempre un subgrupo normal propio.
17. Utilizando Ejerc. 16, hallar todos los grupos abstractos posibles de órdenes a) seis; b) diez; c) catorce; d) quince.
18. Utilizando el análisis del Ejerc. 16, demostrar que hay sólo dos grupos de orden igual al cuadrado de un número primo.

12. Homomorfismos

Una transformación uniforme de un grupo G en un grupo G' puede conservar la multiplicación sin ser biunívoca (esto es, sin ser un isomorfismo).

Así, consideremos una correspondencia entre el grupo simétrico de grado n y el grupo de ± 1 con multiplicación, que lleve las sustituciones pares sobre $+1$ y las impares sobre -1. Según (12), esta correspondencia conserva los productos.

Consideremos ahora la correspondencia entre los elementos a de un grupo abstracto y los automorfismos interiores T_a que ellos determinan. Por la demostración del teorema 22, $T_a T_b = T_{ab}$, luego se conservan los productos. Pero, como se ve en el grupo de simetrías del cuadrado, la correspondencia $a \to T_a$ no es en general biunívoca. (R^2 e I inducen el mismo automorfismo interno.)

Consideremos la correspondencia $n \to i^n$, en la que $i = \sqrt{-1}$, entre el grupo aditivo de los enteros y el multiplicativo de las raíces cuartas de la unidad. También se conservan las operaciones del grupo $i^{m+n} = i^m i^n$, pero la correspondencia es pluriunívoca.

Estos y otros ejemplos nos llevan a la siguiente

DEFINICIÓN. *Se llama* homomorfismo *de un grupo* G *sobre un grupo* G', *a una transformación uniforme* $x \to x'$ *que transporta al conjunto de elementos de* G *sobre la totalidad de* G', *de tal modo, que* $(xy)' = x'y'$ *para cualesquiera* x *e* y *en* G.

TEOREMA 24. *Cualquier homomorfismo* $G \to G'$ *lleva la identidad de* G *sobre la de* G', *y elementos inversos sobre elementos inversos.*

Por hipótesis, la imagen e' de e satisface a $e'a' = (ea)' = a'$ para toda $a' \in G'$; de aquí que e' es la identidad (a la izquierda) de G'. Ahora bien, $aa^{-1} = e$ se transporta sobre $a'(a^{-1})' = e'$. Pero e' es la identidad de G'; luego la imagen $(a^{-1})'$ de la inversa de a es la inversa $(a')^{-1}$ de la imagen de a.

COROLARIO. *Todo grupo homomorfo de un grupo cíclico es cíclico.*

Por el teorema 24, $(a^m)' = (a')^m$, ya sea m positivo, cero o negativo. De aquí que si las potencias a^m agotan G, las potencias $(a')^m = (a^m)'$ de a' agotarán G'.

TEOREMA 25. *El conjunto N de todos los elementos de G a los que corresponde la identidad e' de G', en un homomorfismo de G sobre G', forman un subgrupo normal de G.*

Este conjunto N se llama *núcleo* del homomorfismo.

Como $e \to e'$, N es no vacío. Además, por el teorema 24 y la hipótesis, vemos que $a \to e'$ y $b \to e'$ implican $a^{-1} \to (a')^{-1} = e'^{-1} = e'$ y $ab \to a'b' = e'e' = e'$; luego N es un subgrupo. Finalmente, a en G y x en N implican $a^{-1}xa \to a'^{-1}x'a' = a'^{-1}e'a' = e'$; por lo tanto, N es normal.

EJERCICIOS

1. En el homomorfismo $n \to i^n$, hallar el subgrupo representado sobre la identidad y mostrar que es normal.
2. Mostrar que un grupo cíclico de orden 8 es homomorfo con a) un grupo cíclico de orden 4; b) un grupo cíclico de orden 2.
3. ¿Es un homomorfismo del grupo aditivo de los números reales x la correspondencia que representa $x \to e^{2\pi ix}$ (con $i = \sqrt{-1}$)? Caso afirmativo, ¿quién es G' y cuál es el grupo representado sobre 1?
4. Si G es un grupo de sustituciones entre n letras 1, 2, ..., n, en el cual cada permutación ϕ de G transporta el subconjunto 1, ..., k, sobre sí mismo, mostrar que G es homomorfo con el grupo G' de sustituciones ϕ^* inducidas sobre 1, ..., k.
5. Sean d y d' las dos diagonales de un cuadrado, y sean h y v los ejes. Mostrar que existe un homomorfismo $\phi \to \phi^*$ en el cual cada movimiento ϕ en el grupo del cuadrado induce una permutación ϕ^* sobre d, d', h y v. Mostrar con detalle la correspondencia $\phi \to \phi'$. ¿Cuál es el subgrupo representado sobre I?
6. Si G es homomorfo con G', y G' con G'', demostrar que G es homomorfo con G''.
7. ¿Cuáles de las siguientes correspondencias representan homomórficamente al grupo multiplicativo de todos los números reales no nulos sobre una parte de él mismo? Si la correspondencia es un homomorfismo, identificar el grupo homomorfo G' y el subgrupo representado sobre 1.

   a) $x \to |x|$; b) $x \to 2x$; c) $x \to x^2$; d) $x \to 1/x$;
   e) $x \to -x$; f) $x \to x^3$; g) $x \to -1/x$; h) $x \to \sqrt{x}$.

*13. Grupo cociente

Ahora vamos a ocuparnos en la construcción de réplicas isomorfas de todas las imágenes homomorfas G' de un grupo abstracto dado G.

A tal efecto, sea $x \to x'$ un homomorfismo de G a G', y sea N el subgrupo normal de G representado en la identidad e' de G'. Si a y b son dos elementos cualesquiera de G, podremos escribir $b=at$, así que $b'=a't'$. Pero, por la ley de simplificación, $a't'=a'$ si, y sólo si, $t'=e'$, esto es *(supra)* si, y sólo si, $b=ax$ $(x \in N)$.

LEMA 1. *Dos elementos de* G *tienen la misma imagen en* G' *si, y sólo si, están en el mismo cogrupo* $Nx=xN$ *de* N.

Se establece así una correspondencia biunívoca entre los elementos de G' y los cogrupos de N en G. Por lo tanto, el orden de G' es el número de cogrupos (o «índice») de N en G.

LEMA 2. *Sean* x' *e* y' *dos elementos de* G'. *El producto* x'y' *puede establecerse como sigue: sean* Nx *y* Ny *los cogrupos correspondientes con* x' *e* y' *respectivamente: entonces,* x'y' *corresponde al (único) cogrupo de* N *que contiene al conjunto* NxNy *de todos los productos* uv $(u \in Nx,\ v \in Ny)$.

Demostración. Si $u=ax$, $v=by$ $(a,\ b \in N)$, se tendrá:

$$(uv)'=a'x'b'y'=e'x'e'y'=x'y'.$$

De este modo, G' está determinado, salvo isomorfismos, por G y N: es isomorfo con el sistema de los cogrupos de N en G, multiplicados por la regla de que el «producto» $Nx \circ Ny$ de dos cogrupos es el (único) cogrupo que contiene todos los productos uv $(u \in Nx,\ v \in Ny)$.

Podemos ilustrar las precedentes consideraciones atendiendo al homomorfismo entre el grupo G de las simetrías del cuadrado y el grupo del cuadrilátero G': $[e, a, b, c]$ (§ 9), en el que $[I, R^2] \to e$, $[R, R^3] \to a$, $[H, V] \to b$, $[D, D'] \to c$. (Comprobar sobre la tabla del grupo que esta correspondencia es un homomorfismo.) Los antecedentes de e forman el subgrupo normal $[I, R^2]$, y los de cada uno de los otros elementos son los cogrupos de $[I, R^2]$. Finalmente, podemos deducir la sencilla regla $ab=c$ calculando los productos

$[RH, RV, R^3H, R^3V]$, los cuales están en (o mejor, forman) el cogrupo $[D, D']$, antecedente de c.

Recíprocamente, sea N un subgrupo normal dado en G, no asociado «a priori» con ningún homomorfismo. Se podrá construir un G' homomorfo de G, utilizando N como sigue:

Definiremos los elementos de G' como los distintos cogrupos Nx de N (*). El producto $[Nx]\circ[Ny]$ de dos cogrupos de N es un cogrupo (veremos que único) que contiene al conjunto $NxNy$ de todos los productos uv ($u \in Nx$, $v \in Ny$). Si $u=ax$ y $v=by$, con a y b en N, será $uv=axby=ab'xy$, donde $b'=xbx^{-1}$ es también de N, por ser N normal. Por lo tanto, $N(xy)$ es un cogrupo que contiene a $NxNy$; por otra parte, como los diversos cogrupos no tienen elementos comunes, y el conjunto $NxNy$ no es vacío, es imposible que dos cogrupos distintos comprendan a este conjunto.

De este modo hemos definido una operación binaria uniforme entre los elementos de G' (*alias* cogrupos de G), que puede formularse así:

$$[Nx]\circ[Ny]=N(xy). \tag{17}$$

En palabras: el producto de dos cogrupos se establece multiplicando dos elementos cualesquiera «representantes» de ellos y formando el cogrupo que contiene al producto xy de tales elementos. El producto $N[e]\circ N[y]=N[ey]=Ny$, por (17), así que el cogrupo $N=Ne$ es la identidad a la izquierda para el sistema de cogrupos. Además, $([Nx]\circ[Ny])\circ[Nz]$ y $[Nx]\circ([Ny]\circ[Nz])$ contienen a $(xy)z=x(yz)$, así que la multiplicación de cogrupos es asociativa. Finalmente, $[Nx^{-1}]\circ[Nx]$ contiene a $x^{-1}x=e$, luego debe ser el cogrupo $Ne=N$; por lo tanto, existe el inverso a la izquierda del cogrupo. Estos resultados, con el Teor. 4, prueban lo siguiente:

LEMA 3. *Los cogrupos de un subgrupo normal* N *de* G *forman un grupo multiplicativo.*

DEFINICIÓN. *El grupo de los cogrupos de* N *es llamado el grupo cociente (o bien, grupo factor) de* G *por* N, *y se indica por* G/N (**).

(*) No es ilícito tomar los conjuntos como elementos; así se habla de un «regimiento» indicando cierto conjunto de hombres, o de una «molécula» indicando cierto número de átomos.

(**) Si G es un grupo abeliano en que la operación binaria se denota por «+», cualquier subgrupo de N es normal en G. Y el grupo cociente se llama entonces grupo diferencia, y se indica por $G - N$.

La correspondencia $x \to Nx$ es, por (17), un homomorfismo de G a G/N, y en este homomorfismo los elementos de N son, exactamente, los elementos representados en la identidad de G/N.

Recíprocamente, habíamos visto (a continuación del Lema 2) que para cualquier homomorfismo de G a G', en el cual el subgrupo normal representado sobre la identidad sea N, la imagen G' es isomorfo con el grupo cociente G/N. En conclusión, obtenemos

TEOREMA 26. *Las imágenes homomorfas de un grupo abstracto dado* G *son los grupos cocientes* G/N *por sus diferentes subgrupos normales, con la multiplicación de cogrupos definida por (17).*

Nota. La precedente «construcción» de los grupos cocientes a partir de los grupos y sus subgrupos normales es análoga a la construcción de un dominio de enteros «mód. n» partiendo del dominio de integridad de todos los enteros (Cap. I, §§ 9-10). Los cogrupos de N son los análogos de las clases residuales mód. n, y la relación $x \equiv y$ (mód. n) puede definirse paralelamente así: $x \equiv y$ (mód. N) cuando $xy^{-1} \in N$; esto es equivalente a decir que x e y están en un mismo cogrupo de N (ver Ejerc. 6 de § 9).

EJERCICIOS

1. Enumerar todos los grupos abstractos que son imagen homomorfa del grupo de las simetrías del cuadrado.
2. Hacer lo mismo para el grupo del hexágono regular.
3. Demostrar que el centro Z (§ 8, Ejerc. 7) de cualquier grupo G es un subgrupo normal de G, y que G/Z es isomorfo con el grupo de automorfismos internos de G.
4. Demostrar que en § 9, Ejerc. 6, $x \equiv y$ (mód. S) implica $ax \equiv ay$ (mód. S) para todo a si, y sólo si, S es un subgrupo normal.
5. Si G es el grupo de todos los números racionales de la forma $2^k3^m5^n$, con exponentes k, m y n enteros, mientras S es el subgrupo multiplicativo de todos los números 2^k, describir: a) los cogrupos de S; b) G/S.
6. Sea $G \to G'$ un homomorfismo. Mostrar que el conjunto de todos los antecedentes de cualquier subgrupo S' de G' es un subgrupo S de G, y que si S' es normal, también lo será S.
7. Si S es un subgrupo y N un subgrupo normal de un grupo G, si $S \frown N = e$ y $S \smile N = G$, demostrar que G/N es isomorfo con S.
*8. Si G es un grupo, se llaman conmutadores a los elementos de la forma $x^{-1}y^{-1}xy$. Demostrar que el conjunto C de todos los productos de tales conmutadores forma un subgrupo normal de G.
*9. En Ejerc. 8 demostrar que G/C es abeliano. Finalmente, si N es un subgrupo normal de G y G/N es abeliano, demostrar que $N \geqslant C$.

*10. Dos subgrupos S y T de un grupo G se llaman *conjugados* si $a^{-1}Sa=T$ para algún $a \in G$. Demostrar que la intersección de cualquier subgrupo S de G con sus conjugados es un subgrupo normal de G.

*11. Sea G cualquier grupo, S cualquier subgrupo de G. Para cualquier $a \in G$ sea T_a la permutación $(Sx) \to (Sxa)$ sobre los cogrupos a la derecha Sx de S. Demostrar:

a) La correspondencia $a \to T_a$ es un homomorfismo;

b) El subgrupo transportado sobre la identidad es el subgrupo normal del Ejerc. 10.

14. Equivalencia abstracta y relación de congruencia

Al definir la relación $a \equiv b$ (mód. n) entre enteros, así como al establecer, para introducir los números racionales, que la congruencia entre los pares de números $(a, b) \equiv (a', b')$ significaba, por definición, $ab' = a'b$, y en otras ocasiones, hemos afirmado que cualquier relación reflexiva, simétrica y transitiva podía considerarse como una especie de igualdad. Ahora vamos a formular el significado de esta afirmación.

Convendremos en que una relación R que tenga las propiedades reflexiva, simétrica y transitiva,

$$aRa, \quad aRb \text{ implica } bRa, \quad aRb \text{ y } bRc \text{ implican } aRc,$$

para todos los elementos a, b, c, de un conjunto S, será llamada una *relación de equivalencia* abstracta en S. Si consideramos ciertos *subconjuntos* de S como elementos (tal es el caso de los cogrupos en § 13), tal relación de equivalencia coincide con la igualdad ordinaria. En efecto, si a es cualquier elemento de S, designemos por $R(a)$ el conjunto de todos los elementos de S equivalentes al a; $b \in R(a)$ si, y sólo si, bRa. Estos *R-subconjuntos* tienen varias propiedades sencillas.

LEMA 1. aRb *implica* R(a) = R(b), *y recíprocamente*.

Demostración. Supongamos primero que aRb, y sea c cualquier elemento de $R(a)$. Entonces cRa, por definición; luego, por la ley transitiva, cRb, lo cual significa que $c \in R(b)$. Viceversa, como por la ley simétrica bRa, $c \in R(b)$ implicará $c \in R(a)$, lo cual significa que los dos conjuntos $R(a)$ y $R(b)$ tienen los mismos elementos y por consiguiente son iguales.

Supongamos ahora que $R(a)=R(b)$. Por la ley reflexiva, bRb, así que $b \in R(b)$. Como $R(a)=R(b)$, esto implica $b \in R(a)$ y, por lo tanto, aRb. Esto completa la demostración.

En el caso particular de que R sea la relación de congruencia mód. n entre enteros, el conjunto $R(a)$ determinado por un entero a es, simplemente, la clase residual que contiene a a. El Lema 1 afirma en este caso particular que $a \equiv b$ (mód. n) si, y sólo si, a y b están en la misma clase residual mód. n (cfr. Cap. I, § 10). Otros ejemplos ilustrativos se dan en los ejercicios.

Además, las clases residuales módulo n dividen el conjunto J de todos los enteros en ciertos subconjuntos sin elementos comunes, por lo que se puede decir que determinan una «partición» de J. En general, una partición π de un conjunto S es cualquier colección de subconjuntos A, B, C, ... de S tales, que cada elemento de S se encuentre en uno, y sólo en uno, de los subconjuntos de la colección. Los R-subconjuntos proporcionan siempre una tal partición.

LEMA 2. *Dos* R-*subconjuntos son idénticos o no tienen ningún elemento común, y la colección de todos los* R-*subconjuntos es una partición de* S.

Demostración. Si $R(a)$ y $R(b)$ contuviesen algún elemento c común, sería cRa y cRb, y entonces, por las leyes simétrica y transitiva, aRb. Por el Lema 1 esto implica $R(a)=R(b)$. En consecuencia, si $R(a) \neq R(b)$, las dos clases no pueden tener ningún elemento común. Finalmente, cualquier elemento c de S pertenece al R-subconjunto $R(c)$, ya que, por la ley reflexiva, cRc, y por ende $c \in R(c)$.

La recíproca de los Lemas 1 y 2 es inmediata. Si un conjunto S se divide por una partición π en clases A, B, C, ..., la relación aRb significará, por definición, que los elementos a y b de S pertenecen a un mismo subconjunto en esta partición, y así tenemos una relación de equivalencia abstracta en S. Además, el R-subconjunto $R(a)$ determinado por cada elemento a según esta relación, es exactamente la clase de la partición π que contiene a a. Todas estas conclusiones pueden resumirse en lo que sigue:

TEOREMA 27. *Toda relación* R *de equivalencia abstracta en un conjunto* S *determina una partición* π *de* S *en* R-*clases sin elemen-*

os comunes, y, recíprocamente, cada partición de S *proporciona* *ına relación* R. *Para un* S *dado existe así una correspondencia bi-* *nívoca* R$\leftrightarrow\pi$, *entre las relaciones* R *de equivalencia abstracta en* *S y las particiones* π *de* S, *de tal modo, que dos elementos* a *y* b *le* S *pertenecen a la misma subclase de la partición* π *si, y sólo si,* ıRb.

Discutiendo las condiciones para que una relación de igualdad uese admisible (Cap. I, § 11), fué exigida una cierta «propiedad de ustitución» relativa a las operaciones binarias. Mediante las relaiones de equivalencia R y las operaciones binarias $a \circ b = c$ en S, sta propiedad toma la forma

(18) aRa' y bRb' implican $(a \circ b)R(a' \circ b')$.

Esta condición tiene asimismo un contenido teórico definido.

En efecto, sea R cualquier relación abstracta de equivalencia en S, y sea π la correspondiente partición en R-subclases A, B, C, ... Exactamente como en los cogrupos, podemos considerar estas R-subclases como los elementos de un nuevo conjunto Σ. Y, exactamente como en el grupo cociente (o como en las clases residuales mód. n), podemos intentar definir una operación binaria en Σ a partir de la de S,

(19) $A \circ B = C$ en Σ

si, y sólo si, $a \in A$ y $b \in B$ implica $(a \circ b) \in C$ en S.

La propiedad (18) asegura que si a y a' están ambas en una R-subclase A (es decir, si aRa') y si b y b' están en una R-subclase B, entonces $a \circ b$ y $a' \circ b'$ pertenecen ambas a una misma R-subclase. Esta R-subclase resultante C está unívocamente determinada por A y B, y es el «producto» $A \circ B$ en el sentido de (19). En otras palabras, el principio de sustitución (18) equivale a afirmar que la definición (19) da una operación binaria (univalente) en las R-subclases (es decir, en Σ). Esto demuestra :

Teorema 28. *Dada una relación de equivalencia abstracta en un conjunto* S, *cualquier operación binaria definida en* S *que tenga la propiedad de sustitución (18), determina una operación binaria univalente entre los* R*-subconjuntos de* S, *definida por (19).*

Este resultado se aplica a sistemas como los anillos conmutativos, en los que hay dos operaciones binarias. Y puede también

generalizarse a operaciones no binarias, y aplicarse así a cualquier «sistema algebraico». Las relaciones de equivalencia de tales sistemas, que cumplen un principio de sustitución semejante al (18), pueden llamarse «relaciones abstractas de congruencia».

EJERCICIOS

1. ¿Cuáles de las siguientes relaciones R son relaciones de equivalencia? En el caso de que lo sean, describir las R-subclases:
   a) G es un grupo, S un subgrupo y aRb significa $a^{-1}b \in S$.
   b) G, S como en a); aRb significa $ba^{-1} \in S$.
   c) J es el dominio de los enteros; aRb significa que $a - b$ es primo.
   d) J como en c), aRb significa que $a - b$ es par.
   e) J como en c), aRb significa que $a - b$ es impar.
2. Sea G un grupo de permutaciones de las letras $x_1, \ldots, x_n$; la relación $x_i R x_j$ significa que $x_i\phi = x_j$ para alguna $\phi \in G$. ¿Es R una relación de equivalencia? ¿Qué efecto tiene G sobre cada R-subclase?
3. Sea G el conjunto de las transformaciones $(x, y) \to (x+a, y)$ del plano. Pondremos $(x, y)R(x', y')$ para significar que $(x, y)\phi = (x', y')$ para algún $\phi \in G$. ¿Cuántos subclases R hay en este caso?
4. Siendo a, b números reales, indiquemos por aRb que $a - b$ es un entero múltiplo de 360:
   a) ¿Es R una relación de equivalencia?
   b) ¿Es una relación de congruencia para la adición?
   c) ¿Lo es para la multiplicación?
   d) ¿Qué implica esto respecto a la adición y multiplicación de *ángulos*?
5. a) Enunciar concretamente lo que significa el Teorema 28 cuando se aplica a un anillo conmutativo A.
   b) Demostrar que si R es una relación de congruencia sobre un anillo conmutativo, las R-subclases forman otro anillo conmutativo si la adición y multiplicación se definen por (19). Ilustrarlo por el caso $A = J$.
6. En el Ejerc. 1 a), mostrar que parte de la regla de sustitución (18) vale para cualquier S y que la otra parte vale si, y sólo si, S es normal.

CAPÍTULO VII

Vectores y espacios vectoriales

1. Ejemplo inicial

Aparecen en Física ciertas magnitudes, a las que usualmente se llama vectores, que no son meramente números, sino que además de un valor numérico tienen una dirección. El efecto de un desplazamiento en el plano, por ejemplo, depende, no sólo de la longitud, sino también de la dirección del desplazamiento. Puede representarse convenientemente por una flecha α de longitud y dirección apropiadas (fig. 1). El efecto combinado de dos desplazamientos de este tipo α y β ejecutados uno después de otro, es un tercer desplazamiento «total» γ. Si β es aplicado después de α, colocando el origen de la flecha β en el extremo de la α, el desplazamiento combinado $\gamma = \alpha + \beta$ es la flecha que va del origen de α al extremo de β. Es, pues, la diagonal del paralelogramo de lados α y β. Este modo de hallar $\alpha + \beta$ es la llamada *regla del paralelogramo*.

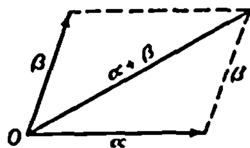


Figura 1

Un desplazamiento α puede triplicarse, obteniéndose así un desplazamiento 3α, o ser reducido a su mitad, con lo que resulta un desplazamiento $\alpha/2$. También se puede representar un desplazamiento múltiplo negativo del α, como -2α, formando un desplazamiento de doble longitud que α y en sentido opuesto a α. En general, α puede ser multiplicado por cualquier número real c para formar un nuevo desplazamiento $c\alpha$. Si c es positivo, $c\alpha$ es del

mismo sentido que α y de longitud c veces mayor; mientras que si c es negativo, el sentido ha de invertirse. Los números c se llaman «escalares» y el producto $c\alpha$ se llama producto escalar (*).

Las fuerzas que actúan sobre un punto de un plano, las velocidades en un plano, los momentos, las aceleraciones, etc., tienen representaciones análogas por medio de vectores, y en todos estos casos, la regla del paralelogramo para la adición de vectores y la multiplicación por escalares (reales) tienen el mismo significado que en los desplazamientos. Ello es un ejemplo de cómo diversas concepciones físicas pueden tener la misma representación matemática.

La Geometría analítica nos sugiere la representación de los vectores en un plano, por pares de números reales. Podemos representar cada vector por una flecha con origen en (0, 0) y extremo en un punto (a_1, a_2). Entonces, la suma de vectores y los productos por escalares pueden calcularse con estas reglas:

(1) $$(a_1, a_2)+(b_1, b_2)=(a_1+b_1, a_2+b_2),$$

(2) $$c(a_1, a_2)=(ca_1, ca_2).$$

Partiendo de tales reglas podemos obtener fácilmente muchas leyes de cálculo algebraico vectorial (**), tales como:

(3) $$\alpha+\beta=\beta+\alpha, \qquad \alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma,$$

(4) $$c(\alpha+\beta)=c\alpha+c\beta, \qquad 1\cdot\alpha=\alpha,$$

y así sucesivamente. Muchas de ellas (como la ley conmutativa de la suma vectorial) corresponden a principios geométricos bien conocidos.

Las operaciones con vectores pueden utilizarse ya para formular algunas sencillas propiedades geométricas. Por ejemplo, el punto medio de la línea que une el extremo del vector $\alpha=(a_1, a_2)$ con el del $\beta=(b_1, b_2)$ viene dado por las fórmulas $[(a_1+b_1)/2, (a_2+b_2)/2]$, o sea, por el vector semisuma $(1/2)(\alpha+\beta)$. Este vector da también el centro de gravedad de α y β. Una lista completa de las leyes de álgebra vectorial será dada en el § 3; vamos antes a presentar otros ejemplos de vectores.

(*) La denominación «producto escalar» tiene corrientemente, en castellano, otra significación, equivalente a la de «producto interno» (cfr. § 7). Ponemos al lector en guardia contra la posible confusión entre ambos conceptos: «producto (por un) escalar» y «producto escalar (de dos vectores)». (N. del T.)

(**) Los vectores se designarán con minúsculas griegas, α, β, ..., y los escalares con minúsculas latinas, a, b, ...

EJERCICIOS

1. Demostrar las leyes (3) y (4) del álgebra vectorial, utilizando las reglas (1) y (2).
2. Ilustrar la ley distributiva (4) por un diagrama.
3. Mostrar que los vectores del plano forman un grupo bajo la adición.
4. Demostrar que cualquier vector α en el plano puede representarse unívocamente como una suma $\alpha=\beta+\gamma$, donde β es un vector sobre el eje x, y γ un vector sobre el eje y.

2. Generalizaciones

El ejemplo que se acaba de dar puede generalizarse de dos maneras. En primer lugar, el número de *dimensiones* (que era dos en el § 1) puede ser arbitrario. El primer apoyo intuitivo para esto lo encontramos en la posibilidad de tratar las fuerzas y desplazamientos en el espacio, del mismo modo que hemos tratado, en el párrafo anterior, las fuerzas y desplazamientos en el plano. La única diferencia es que, en el caso del espacio, los vectores tienen *tres* componentes (x_1, x_2, x_3) en vez de dos.

Además, se demuestra en Estática que las fuerzas que actúan sobre un cuerpo rígido pueden reducirse a una fuerza y un par cuyo eje tenga la misma dirección que la fuerza, con lo que dicho sistema se reduce a *seis* componentes respecto de ejes perpendiculares. La suma de dos o más fuerzas puede calcularse componente a componente, y la multiplicación por escalares (números reales) tiene la misma significación que antes explicamos.

Mas generalmente, para cualquier número positivo n, los conjuntos $\alpha=(a_1, \ldots, a_n)$ de números reales forman un *espacio n-dimensional* de vectores en el que puede considerarse una geometría de n dimensiones. Así, líneas rectas son los conjuntos de elementos de la forma $\alpha+x\beta$ (α, $\beta\neq 0$ fijos, x variable); el centro de gravedad de $\alpha_1, \ldots, \alpha_n$ puede definirse como $(1/n)(\alpha_1+\ldots+\alpha_n)$ y así sucesivamente (esta idea será desarrollada en el Cap. IX). Para obtener una teoría geométrica completa necesitamos solamente introducir el concepto de distancia, como lo haremos en el § 8.

De hecho, los espacios físicos de elevado número de dimensiones (cada coordenada corresponde a un grado de libertad) han sido muy utilizados en Dinámica. A los matemáticos puros les parece muy modesto usar solamente cuatro dimensiones, como en relatividad (tres componentes de espacio y una de tiempo). Para ilus-

trar un poco más este punto, vamos a construir ahora un espacio vectorial de infinitas dimensiones, usado comúnmente en Análisis matemático.

Indiquemos con S el conjunto de todas las funciones $f(x)$ de una variable real x, uniformes y continuas en el intervalo $0 \leqslant x \leqslant 1$. Dos de tales funciones, $f(x)$ y $g(x)$, pueden sumarse para dar otra función, $h(x) = f(x) + g(x)$, y el producto «escalar» de $f(x)$ por una constante real c es también una función $cf(x)$. Estas funciones no pueden representarse por flechas, pero son susceptibles de las operaciones de adición y multiplicación escalar con las mismas propiedades formales algebraicas que las de nuestros anteriores ejemplos. Puede considerarse que en este conjunto S los «vectores» tienen una «componente» en cada punto x de la línea $0 \leqslant x \leqslant 1$ (esta componente es el valor de la función).

Una segunda posibilidad de generalización puede deducirse observando que, en lo que se refiere a propiedades algebraicas, las componentes de los vectores y los escalares no necesitan ser números reales, sino que pueden ser elementos de cualquier campo. Así, los vectores de componentes *complejos* se usan constantemente en la teoría de circuitos eléctricos y en el electromagnetismo, mientras que nosotros, en el capítulo XIV, emplearemos los vectores con *escalares* racionales para desarrollar la teoría general de números algebraicos.

Para concretar, definiremos el sistema $V_n(F)$, para cualquier entero positivo n y cualquier campo F, como el conjunto de todas las n-plas $\alpha = (a_1, \ldots, a_n)$, $\beta = (b_1, \ldots, b_n)$, ..., con componentes a_i y b_i en F. Las dos operaciones vectoriales de «adición» y «multiplicación escalar» (esto es, por elementos c de F) se definirán como sigue:

(5) $$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n),$$

(6) $$c(a_1, \ldots, a_n) = (ca_1, \ldots, ca_n).$$

Veremos en § 3 que además de las leyes (3) y (4), mencionadas en § 1, hay otras muchas que se cumplen en $V_n(F)$.

EJERCICIOS

1. Sea $\alpha = (1, 1, 0)$, $\beta = (-1/2, 0, 2/3)$, $\gamma = (0, 1/4, 2)$. Calcular:
   a) $\alpha + 2\beta + 3\gamma$; b) $3(\alpha + \beta) - 2(\beta + \gamma)$;
   c) ¿Cuál es el centro de gravedad de α, β, γ?
   d) Resolver $6\beta + 5\xi = \alpha$.

2. Sea $\alpha=(1, i, 0)$, $\beta=(0, 1-i, 2i)$, $\gamma=(1, 2-i, 1)$. Calcular:
   a) $2\alpha-i\beta$; b) $i\alpha+(1+i)\beta-(i+3)\gamma$;
   c) Resolver $\alpha-i\xi=\beta$.
3. Dividir el segmento rectilíneo $\overrightarrow{\alpha\beta}$ en la razón 2:1 en Ejerc. 1 y Ejerc. 2.
4. En el Ejerc. 2 ¿puede dividirse el segmento $\overrightarrow{\alpha\beta}$ en la razón 1:2i? Explicarlo.
5. Sea $V_n(J_3)$ el conjunto de vectores con n componentes en el campo de enteros mód. 3. a) ¿Cuántos vectores hay en $V_n(J_3)$? b) ¿Qué puede decirse sobre el vector $\alpha+\alpha+\alpha$ en $V_n(J_3)$?
6. ¿Puede el lector definir un «punto medio» entre dos puntos arbitrarios en $V_n(J_3)$? ¿Y un centro de gravedad para tres — para cuatro — puntos arbitrarios? Buscar ejemplos numéricos.

3. Espacios vectoriales y subespacios

Las propiedades algebraicas de los vectores se resumen por entero en la siguiente

DEFINICIÓN. *Se llama espacio vectorial (y también «espacio lineal») V, sobre un campo F, a un conjunto de elementos, llamados vectores, tales, que dos cualesquiera de ellos α y β determinan unívocamente un vector suma α+β, y cada vector α de V y cada escalar c de F determinan un producto escalar c·α en V, con las propiedades siguientes:*

(7) V *es un grupo abeliano aditivo.*

(8) $$c\cdot(\alpha+\beta)=c\cdot\alpha+c\cdot\beta, \qquad (c+c')\cdot\alpha=c\cdot\alpha+c'\cdot\alpha$$
(leyes distributivas)

(9) $$(cc')\cdot\alpha=c\cdot(c'\cdot\alpha), \qquad 1\cdot\alpha=\alpha.$$

(Las reglas (8) y (9) son válidas para todos los vectores α y β y todos los escalares c y c'.)

EJEMPLOS. La adición y la multiplicación escalar de n-plas, tal como fué definida por (5) y (6), satisface estas leyes formales; por ejemplo, las definiciones (5) y (6) reducen cada una de las leyes distributivas (8) a la correspondiente ley distributiva para cada componente, que ciertamente es válida. Por lo tanto, el espacio $V_n(F)$ de tales n-plas es un espacio vectorial sobre F, en el sentido de nuestra definición.

Consideremos ahora las funciones f cuyo dominio de definición es un conjunto cualquiera S, siendo la resultante el campo F; así

que f hace corresponder a cada x en S un valor $f(x)$ en F. El conjunto de tales funciones f constituye un espacio vectorial sobre F, si la suma $h=f+g$ y el producto escalar $h'=c\cdot f$ son las funciones definidas para cada x en S, por las igualdades $h(x)=f(x)+g(x)$ y $h'(x)=c\cdot f(x)$.

De acuerdo con la acostumbrada notación aditiva para expresar la operación del grupo, representaremos por O el elemento de identidad del grupo: es el único vector «nulo» o «cero» que satisface

$$(10) \qquad \xi+O=O+\xi=\xi, \text{ para todo } \xi.$$

No debe confundirse el vector O con el escalar 0. Sin embargo, ambos están relacionados por una identidad. En efecto, las dos leyes distributivas dan, para todo c y ξ,

$$c\xi+0\xi=(c+0)\xi=c\xi=c\xi+O,$$
$$c\xi+cO=c(\xi+O)=c\cdot\xi=c\xi+O.$$

Ahora, cancelando $c\xi$ en ambos miembros, obtenemos las dos leyes

$$(11) \qquad 0\xi=O \text{ para todo } \xi, \qquad cO=O \text{ para todo } c.$$

El múltiplo escalar $(-1)\xi$ actúa como el opuesto de cualquier vector ξ del grupo, ya que

$$\xi+(-1)\xi=1\cdot\xi+(-1)\xi=[1+(-1)]\xi=0\xi=O;$$

por lo tanto,

(12) el *inverso* aditivo de cada vector ξ es $(-1)\xi$.

Se deduce de (11) y (12) que el subgrupo de las «potencias» (Cap. VI, § 6) de un vector ξ está ahora constituído por los *múltiplos enteros* $n\xi$ de ξ.

En el espacio vectorial tridimensional ordinario, $V_3(R^*)$, los vectores situados sobre un determinado plano fijo que pase por el origen, forman por sí mismos un espacio vectorial bidimensional, que es parte del espacio entero. Análogamente, el conjunto de todos los vectores situados sobre una recta que pasa por el origen es cerrado para las operaciones de adición y de multiplicación por escalares, y, por tanto, este conjunto es también un «subespacio» de $V_3(R^*)$.

DEFINICIÓN. *Un subespacio* S *de un espacio vectorial* V *es un subconjunto de* V *que es también un espacio vectorial con respecto a las mismas operaciones de adición y multiplicación escalar definidas en* V.

La condición necesaria y suficiente para que S sea un subespacio es que la suma de dos vectores cualesquiera de S esté en S, y que el producto de cualquier vector de S por un escalar esté en S. Esta proposición puede sacarse con facilidad de la definición. La analogía con las definiciones anteriores de subcampo y subgrupo es manifiesta. Geométricamente, un «subespacio» es, simplemente, un subespacio lineal (recta, plano, etc.) que contiene al origen O.

Por ejemplo, los vectores de la forma $(0, x_2, 0, x_4)$ constituyen un subespacio de $V_4(F)$ para cualquier campo F. Notemos que el vector nulo O es por sí solo un subespacio de cualquier espacio vectorial.

Asimismo, el conjunto de los polinomios de grado igual o menor que 7 es un subespacio del espacio vectorial de todos los polinomios, sea real o no el campo base. Análogamente, el conjunto de todas las funciones continuas $f(x)$ definidas para $0 \leqslant x \leqslant 1$ es un subespacio del espacio lineal de todas las funciones definidas en el mismo dominio.

Dados m vectores $\xi_1, \xi_2, \ldots, \xi_m$ de un espacio vectorial V, *el conjunto de todas las combinaciones lineales*

$$c_1\xi_1 + c_2\xi_2 + \ldots + c_m\xi_m \qquad \text{(cada } c_i \text{ es un escalar)}$$

de las ξ_i *es un subespacio.* Esto se desprende de las identidades

$$(13) \quad (c_1\xi_1 + \ldots + c_m\xi_m) + (c_1'\xi_1 + \ldots + c_m'\xi_m) = \\ = (c_1 + c_1')\xi_1 + \ldots + (c_m + c_m')\xi_m,$$

$$(14) \qquad c'(c_1\xi_1 + \ldots + c_m\xi_m) = (c'c_1)\xi_1 + \ldots + (c'c_m)\xi_m,$$

válidas para todos los vectores ξ_i y para todos los escalares c_i, c_i' y c'. Así tenemos:

TEOREMA 1. *El conjunto de todas las combinaciones lineales de un conjunto de vectores de un espacio* V *es un subespacio de* V.

Este espacio es evidentemente el menor subespacio que contiene los vectores dados; por eso se le llama subespacio engendrado por ellos. El subespacio engendrado por un solo vector $\xi \neq 0$ es el conjunto S de todos sus múltiplos escalares $c\xi$; geométricamente es,

sencillamente, la recta que pasa por el origen y ξ_1. De modo análogo, el subespacio engendrado por dos vectores no colineales ξ_1 y ξ_2 viene a ser el plano que pasa por el origen, por ξ_1 y por ξ_2.

TEOREMA 2. *La intersección* $S \frown T$ *de dos subespacios cualesquiera de un espacio vectorial* V *es también un subespacio de* V.

La *intersección* de dos subespacios dados S y T se define como el conjunto $S \frown T$ de todos los vectores que pertenecen simultáneamente a S y a T (cfr. Teorema 17 del Cap. VI, sobre intersección de dos subgrupos). Si α y β son dos vectores en estas condiciones, su suma $\alpha+\beta$ debe estar en S (ya que S es un subespacio que contiene α y β) y en T; por lo tanto, estará también en la intersección $S \frown T$. Análogamente, todo múltiplo escalar $c\alpha$ estará en $S \frown T$ siempre que lo esté α.

Dos subespacios S y T de un espacio vectorial determinan un conjunto $S+T$ formado por todas las sumas $\alpha+\beta$ (con α en S y β en T). Por las leyes conmutativa, asociativa y distributiva (3) y (4), este conjunto es también un subespacio, que se llama la *suma lineal* de S y T, en el cual S y T están contenidos, estándolo $S+T$ en cualquier otro subespacio R que contenga a S y T; de aquí que el concepto de suma lineal es análogo al de unión de dos subgrupos (cfr. Cap. VI, §8). Estas propiedades de $S+T$ pueden expresarse así:

$$(15) \qquad S \leqslant S+T, \qquad T \leqslant S+T;$$
$$S \leqslant R \text{ y } T \leqslant R \text{ implican } S+T \leqslant R.$$

En ellas, $S \leqslant R$ significa que el subespacio S está contenido en el subespacio R.

EJERCICIOS

1. Enunciar tres «leyes distributivas generalizadas» válidas en todo espacio vectorial.
2. En Ejerc. 1, § 2, calcular $7[2(\alpha-3\beta)+(1/3)(3\beta-6\gamma)]-2(\alpha-\gamma)+5\beta+2\alpha$.
3. En Ejerc. 2, § 2, calcular $(1+2i)[(2-i)(2\alpha-3\beta)]-8\alpha-9i\beta$.
4. ¿Cuáles de los siguientes subconjuntos de $V_n(R)$ $(n>2)$ constituyen subespacios? [En lo que sigue, ξ indica $(x_1, \ldots, x_n)$.]
   a) Todos los ξ, con x_1 entero;
   b) Todos los ξ, con $x_2=0$;
   c) Todos los ξ, con x_1 o x_2 nulos;
   d) Todos los ξ tales, que $3x_1+4x_2=1$;
   e) Todos los ξ tales, que $7x_1-x_2=0$.

5. ¿Cuáles de los siguientes conjuntos de funciones reales $f(x)$ definidas en $0 \leqslant x \leqslant 1$ son subespacios del espacio vectorial de todas estas funciones?
   a) Todos los polinomios de grado 4;
   b) Todos los polinomios de grado < 4 (incluso $f(x)=0$);
   c) Todas las funciones f con $2f(0)=f(1)$;
   d) Todas las funciones tales, que $0+f(1)=f(0)+1$;
   e) Todas las funciones positivas;
   f) Todas las funciones que satisfacen la condición $f(x)=f(1-x)$ para todo x.
6. ¿Qué conjuntos de funciones descritas en el Ejercicio 3, § 3, Cap. IV, forman espacios vectoriales sobre el campo R^* de los números reales?
7. Sea S un subespacio de $V_3(R)$ que consiste en todos los vectores de la forma $(0, x_2, x_3)$ y T el subespacio engendrado por $(1, 2, 0)$ y $(3, 1, 2)$. ¿Qué vectores están en $S \frown T$? ¿Cuáles en $S+T$?
8. En $V_3(J_3)$ ¿cuántos vectores son engendrados por $(1, 2, 1)$ y $(2, 1, 1)$? ¿Y por $(1, 2, 1)$ y $(2, 1, 2)$?
9. Mostrar en $V_3(R)$ que el plano $x_3=0$ puede engendrarse por cada uno de los siguientes pares de vectores: $(1, 0, 0)$ y $(1, 1, 0)$; $(2, 2, 0)$ y $(4, 1, 0)$; $(3, 2, 0)$ y $(-3, 2, 0)$.
10. Si S es engendrado por ξ_1 y ξ_2, T por η_1, η_2 y η_3, mostrar que $S+T$ se engendra por ξ_1, ξ_2, η_1, η_2, η_3. Generalizar el resultado.
11. Consideremos un conjunto de cuatro vectores O, α, β y γ, para los que la adición se define por las reglas

$$\alpha+\beta=\gamma, \quad \beta+\gamma=\alpha, \quad \gamma+\alpha=\beta, \quad \alpha+\alpha=\beta+\beta=\gamma+\gamma=O,$$

mientras que los múltiplos escalares por enteros, mód. 2, se definen como $0\alpha=O$, $1\alpha=\alpha$, etc.
   a) Mostrar que esos vectores forman un espacio vectorial sobre J_2;
   b) Mostrar que el grupo aditivo de vectores es isomorfo al grupo del cuadrilátero;
   c) Enumerar los subespacios de este espacio;
   d) Mostrar que este espacio es isomorfo con $V_2(J_2)$.
12. Construir $V_4(J_2)$ y tabular sus subespacios.
13. Demostrar que el conjunto de todas las soluciones $(x_1, \ldots, x_n)$ de un par de ecuaciones lineales homogéneas $a_1x_1+\ldots+a_nx_n=0$, $b_1x_1+\ldots+b_nx_n=0$ es un subespacio de $V_n(F)$, con a_i, b_i, x_i en F.
*14. Demostrar que el postulado $1\cdot\alpha=\alpha$ para un espacio vectorial, no puede demostrarse a partir de los otros postulados. (*Sugerencia:* Construir en el plano un producto pseudoescalar $c \odot \alpha$, proyección de $c\cdot\alpha$ sobre una recta fija.)

4. Independencia lineal

Todavía no hemos dado una definición abstracta de la importante noción geométrica de dimensión de un espacio (o subespacio) vectorial. Más adelante, la definiremos como el menor número de

vectores con los que se pueden engendrar el espacio o subespacio dados.

Así, el espacio ordinario $V_3(R)$ puede engendrarse por los tres vectores (1, 0, 0), (0, 1, 0) y (0, 0, 1), de longitud unidad y situados en los tres ejes coordenados; pero no puede serlo por un conjunto de dos vectores (un conjunto de dos vectores no colineales engendra un plano que pasa por el origen).

Más generalmente, todo $V_n(F)$ puede engendrarse por los n vectores unitarios:

(16) $\epsilon_1 = (1, 0, \ldots, 0)$, $\epsilon_2 = (0, 1, 0, \ldots, 0)$, ..., $\epsilon_n = (0, \ldots, 0, 1)$.

En efecto: se ve claramente que cualquier vector de $V_n(F)$ es una combinación lineal de éstos, ya que

(17) $$(x_1, x_2, \ldots, x_n) = x_1\epsilon_1 + x_2\epsilon_2 + \ldots + x_n\epsilon_n.$$

En § 6 demostraremos que $V_n(F)$ no puede engendrarse por un conjunto de menos de n vectores. Esto justifica el llamar a $V_n(F)$ espacio vectorial n-dimensional sobre el campo F.

Acabamos de ver que los vectores $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ engendran por completo al espacio vectorial $V_n(F)$; tienen, además, la propiedad de que sólo es $x_1\epsilon_1 + x_2\epsilon_2 + \ldots + x_n\epsilon_n = O$ cuando $(x_1, \ldots, x_n) = (0, \ldots, 0)$, o sea, si, y sólo si, $x_1 = x_2 = \ldots = x_n = 0$. Por esto se dice que los vectores unitarios son «linealmente independientes», en el sentido de la siguiente

DEFINICIÓN. *Los vectores $\alpha_1, \ldots, \alpha_m$ son linealmente independientes (sobre* F*) cuando, para los escalares* c_i *en* F,

(18) $$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_m\alpha_m = O \textit{ implica } c_1 = c_2 = \ldots = c_m = 0.$$

Los vectores que no son linealmente independientes se llaman linealmente dependientes.

Una consecuencia de la definición es que cualquier subconjunto de un conjunto linealmente independiente es también linealmente independiente. Ahora bien, la siguiente relación de dependencia es más importante:

TEOREMA 3. *Los vectores no nulos $\alpha_1, \ldots, \alpha_m$ en un espacio* V *son linealmente dependientes cuando alguno de los vectores α_k es una combinación lineal de los precedentes, y sólo en este caso.*

Demostración. En el caso de que el vector α_k sea una combinación lineal de las anteriores, tenemos la relación lineal

$$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_{k-1}\alpha_{k-1} + (-1)\alpha_k = O,$$

con, al menos, un coeficiente, (—1), distinto de cero. De aquí que los vectores son dependientes, por (18).

Recíprocamente, supongamos que los vectores son dependientes, o sea, que $d_1\alpha_1 + d_2\alpha_2 + \ldots + d_m\alpha_m = O$, y tomemos el mayor subíndice k para el cual es $d_k \neq 0$. Podemos despejar α_k y obtenerlo como la combinación lineal

$$\alpha_k = (-d_1/d_k)\alpha_1 + \ldots + (-d_{k-1}/d_k)\alpha_{k-1}.$$

Esto nos da α_k como combinación lineal de los vectores precedentes, excepto en el caso $k=1$. En este caso, $d_1\alpha_1 = O$, con $d \neq 0$, así que $\alpha_1 = O$ contra la hipótesis de que ninguno de los vectores dados es igual a cero. Así queda completada la demostración.

Por ejemplo, los tres vectores $\beta_1 = (2, 0, 0,)$, $\beta_2 = (1, 3, 0)$ y $\beta_3 = (0, -2, 0)$ no pueden engendrar todo el espacio ordinario $V_3(R^*)$, porque están situados en un plano. Podemos expresar esta dependencia lineal por la relación $\beta_1 - 2\beta_2 - 3\beta_3 = O$, o, despejando β_1, por la $\beta_1 = 2\beta_2 + 3\beta_3$. Así, el conjunto $(\beta_1, \beta_2, \beta_3)$ engendra el mismo espacio que su subconjunto (β_2, β_3). Esto nos enseña que

COROLARIO 1. *Un conjunto de vectores es linealmente dependiente cuando contiene un subconjunto propio (es decir, menor) que engendra el mismo subespacio que aquéllos.*

Concretamente, podemos separar del conjunto dado cualquier vector que sea nulo o que sea combinación lineal de los precedentes, y se verá que los vectores restantes engendran el mismo subespacio. Mediante inducción resulta :

COROLARIO 2. *Todo conjunto finito de vectores contiene un subconjunto de vectores linealmente independientes que engendra el mismo subespacio que ellos.*

EJERCICIOS

1. Estudiar la dependencia lineal en los siguientes conjuntos de vectores:
   a) (1, 0, 1), (0, 2, 2), (3, 7, 1) en $V_3(R)$ y $V_3(C)$;
   b) (0, 0, 0), (1, 0, 0), (0, 1, 1) en $V_3(R^*)$;

c) (a, i, $i+1$), (i, -1, $2-i$), (0, 0, 3) en $V_3(C)$;
d) (1, 1, 0), (1, 0, 1), (0, 1, 1) en $V_3(J_2)$ y en $V_3(J_3)$.
En caso de dependencia lineal, seleccionar un subconjunto linealmente independiente que engendre el mismo subespacio.

2. Demostrar que los vectores (a_1, a_2) y (b_1, b_2) en $V_2(F)$ son linealmente dependientes si, y sólo si, $a_1b_2 - a_2b_1 = 0$.
3. Sea $\xi_1 = (1, 1, 1)$, $\xi_2 = (2, 1, 2)$, $\xi_3 = (3, 4, -1)$, $\xi_4 = (4, 6, 7)$. Hallar números c_i no todos nulos, tales que $c_1\xi_1 + c_2\xi_2 + c_3\xi_3 + c_4\xi_4 = O$.
4. Sea $\eta_1 = (1+i, 2i)$, $\eta_2 = (2, -3i)$, $\eta_3 = (2i, 3+4i)$. Hallar todos los números complejos c_i tales, que $c_1\eta_1 + c_2\eta_2 + c_3\eta_3 = O$.
5. Hallar dos vectores que engendren el mismo subespacio que todos los vectores (x_1, x_2, x_3, x_4) que satisfacen a $x_1 + x_2 = x_3 - x_4 = 0$.
6. El mismo Ejerc. 5, para los vectores satisfaciendo a

$$3x_1 - 2x_2 + 4x_3 + x_4 = x_1 + x_2 - 3x_3 - 2x_4 = 0.$$

7. Demostrar que si β no está en el subespacio S, pero sí en el subespacio engendrado por S y α, entonces α está en el subespacio engendrado por S y β.
8. Demostrar que si ξ_1, ξ_2, ξ_3 son independientes en $V_n(R)$, entonces también lo son $\xi_1 + \xi_2$, $\xi_1 + \xi_3$, $\xi_2 + \xi_3$. ¿Es cierto esto en cualquier $V_n(F)$?
9. ¿Cuántos elementos existen en cada subespacio desarrollado por cuatro elementos de $V_7(J_3)$ linealmente independientes? Generalizar este resultado.
10. Definir un «espacio vectorial» sobre un dominio de integridad D. ¿Cuáles de los postulados y teoremas tratados antes fallan en el caso más general?
11. Demostrar que tres vectores con coordenadas racionales son linealmente independientes en $V_3(R^*)$ si, y sólo si, son linealmente independientes en $V_3(R)$. Generalizar este resultado de dos modos distintos.
12. Si los vectores $\alpha_1, \ldots, \alpha_m$ son linealmente independientes, mostrar que el vector β es una combinación lineal de $\alpha_1, \ldots, \alpha_m$ si, y sólo si, los vectores $\alpha_1, \ldots, \alpha_m, \beta$ son linealmente dependientes.

*13. Demostrar que los números reales 1, $\sqrt{2}$ y $\sqrt{5}$ son linealmente independientes sobre el campo de los números racionales.

5. Base de un espacio vectorial

Hay muchos conjuntos de vectores de $V = V_n(F)$ que comparten con los vectores unitarios $\epsilon_1, \ldots, \epsilon_n$ las dos propiedades de ser linealmente independientes y engendrar a V_n. Así los $\alpha_1 = (1, 1, 0)$, $\alpha_2 = (0, 1, 1,)$ y $\alpha_3 = (0, 0, 1)$ tienen estas propiedades en $V_3(F)$. Puesto que

$$O = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = (x_1,\ x_1 + x_2,\ x_2 + x_3)$$

supone que ha de ser $x_1=0$, $x_2=0$ y $x_3=0$: y cada vector (y_1, y_2, y_3) de $V_3(F)$ puede expresarse en la forma

$$(y_1, y_2, y_3)=y_1\alpha_1+(y_2-y_1)\alpha_2+(y_3-y_2+y_1)\alpha_3.$$

Esto nos lleva a formular la siguiente

DEFINICIÓN. *Una base de un espacio vectorial es cualquier subconjunto suyo linealmente independiente, que engendra al espacio entero. Diremos que un espacio vectorial es de dimensión finita cuando tenga una base finita.*

Por ejemplo, consideremos la ecuación diferencial lineal homogénea $\frac{d^2x}{dt^2}-3\frac{dx}{dt}+2x=0$, con coeficientes constantes. Se comprueba fácilmente que la suma $x_1(t)+x_2(t)$ de dos soluciones es una solución, y que el producto de una solución por una constante (real) es una solución. De aquí que el conjunto V de todas las soluciones $x(t)$ de esta ecuación diferencial es un espacio vectorial (es también un subespacio del espacio de todas las funciones $x(t)$ que son derivables dos veces). Si se pudiera encontrar una base para este «espacio solución» V, esto significaría que todas las posibles soluciones podrían expresarse como combinaciones lineales de estas soluciones básicas. En el caso considerado, e^t y e^{2t} son soluciones linealmente independientes, y puede demostrarse que la solución más general es de la forma $x=c_1e^t+c_2e^{2t}$, con las «constantes de integración» c_1 y c_2. El procedimiento que se acaba de bosquejar se emplea actualmente en Mecánica, teoría de las corrientes eléctricas y otros problemas referentes a vibraciones.

También el campo C de todos los números complejos puede considerarse como un espacio vectorial, sobre el campo R^* de los números reales, si de todas las operaciones algebraicas en C atendemos tan sólo a la adición de números complejos y a la multiplicación («escalar») de números complejos por reales. Este espacio tiene dimensión 2, pues 1 e i forman una base; cada uno de estos dos elementos engendra, respectivamente, el subespacio de los números reales y el de los imaginarios puros. Los números $1+i$ y $1-i$ forman otra base (menos cómoda) de C sobre R^*.

Finalmente, el dominio $F[x]$ de todas las formas polinómicas en una indeterminada x sobre un campo F es un espacio lineal so-

es única.

Por analogía con las correspondientes definiciones en los dominios de integridad y grupos, definiremos un *isomorfismo* entre dos espacios vectoriales V y V', sobre un mismo campo F, como una correspondencia biunívoca $\xi \leftrightarrow \xi'$ entre los elementos de V y los de V' tal, que $(\xi + \eta)' = \xi' + \eta'$ y $(c\xi)' = c\xi'$ para todos los vectores ξ, η en V y todos los escalares c en F.

TEOREMA 5. *Todo espacio vectorial V sobre F de dimensión finita es isomorfo con un espacio $V_n(F)$ de n-plas.* (Cfr. § 2.)

Por hipótesis, V tiene una base finita $\alpha_1, \ldots, \alpha_n$, y cada vector ξ de V puede expresarse de modo único como una combinación lineal

(19) $$\xi = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n.$$

bre F, ya que las condiciones características de un espacio vectorial se satisfacen en $F[x]$. La definición de igualdad, aplicada a la ecuación $p(x) = 0$, significa que las potencias $1, x, x^2, x^3, \ldots$ son linealmente independientes sobre F. Deducimos de lo dicho que $F[x]$ tiene una *base* infinita constituída por estas potencias, ya que todo vector (forma polinómica) puede ser expresado como una combinación lineal de un subconjunto finito de esta base.

Según el Corolario 2 del Teorema 3, todo espacio vectorial engendrado por un subconjunto finito, sea o no linealmente independiente, tiene una base finita; vemos así que la condición necesaria y suficiente para que un espacio vectorial tenga dimensión finita es que pueda engendrarse a partir de un conjunto finito.

TEOREMA 4. *La representación de un vector en un espacio V, como combinación lineal de los vectores de una base dada, es única.*

Como la base $\alpha_1, \ldots, \alpha_n$ engendra el espacio, cada vector es una combinación lineal $\xi = x_1\alpha_1 + \ldots + x_n\alpha_n$. Si ξ pudiera representarse por otra combinación lineal tal como $y_1\alpha_1 + \ldots + y_n\alpha_n$, restando y

Los coeficientes escalares x_i que aparecen en esta representación se llaman las *coordenadas* del vector ξ respecto a la base $\alpha_1, \ldots, \alpha_n$. Para multiplicar el vector ξ por el escalar c, podemos usar la ley distributiva :

$$(20) \qquad c\xi = (cx_1)\alpha_1 + (cx_2)\alpha_2 + \ldots + (cx_n)\alpha_n,$$

así que basta multiplicar cada coordenada por el escalar c. De modo semejante puede sumarse ξ con otro vector $\eta = y_1\alpha_1 + \ldots + y_n\alpha_n$, sin más que sumar las coordenadas correspondientes, por la ley asociativa,

$$(21) \qquad \xi + \eta = (x_1 + y_1)\alpha_1 + (x_2 + y_2)\alpha_2 + \ldots + (x_n + y_n)\alpha_n.$$

Este modo de operar con las coordenadas guarda un paralelismo completo con las definiciones (5) y (6) de producto escalar y de suma de las n-plas. Así obtenemos un isomorfismo en la correspondencia biunívoca

$$(22) \qquad \xi = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n \leftrightarrow (x_1, x_2, \ldots, x_n)$$

que transporta cada vector ξ de V_n sobre la n-pla de sus coordenadas.

… l menos teóricamente, el pro- … morfismos) de todos los espa- … El razonamiento puede exten- … caso de infinitas dimensiones.

Todavía no hemos indicado el modo de resolver el problema de expresar un vector dado $\beta = (b_1, \ldots, b_n)$ de $V_n(F)$ como combinación lineal de los elementos de una base que no esté constituída por los vectores unitarios $\epsilon_1, \ldots, \epsilon_n$. Esto no es difícil, sin embargo. Sea la base $\alpha_1 = (a_{11}, \ldots, a_{1n})$, $\alpha_2 = (a_{21}, \ldots, a_{2n})$, …, $\alpha_n = (a_{n1}, \ldots, a_{nn})$. En tal caso, afirmar que $\beta = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n$ es tanto como decir que

$$(b_1, \ldots, b_n) = (x_1a_{11} + \ldots + x_na_{n1}, \ldots, x_1a_{1n} + \ldots + x_na_{nn}).$$

Y así, las coordenadas x_i de β relativas a la base $\alpha_1, \ldots, \alpha_n$ pueden calcularse resolviendo el sistema de ecuaciones lineales simultáneas :

$$a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n = b_1,$$
$$\cdots\cdots\cdots$$
$$a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{nn}x_n = b_n.$$

El método para resolver estas ecuaciones fué explicado en el Capítulo II, § 3. La independencia lineal de $\alpha_1, \ldots, \alpha_n$ asegura en el presente caso la existencia y unicidad de la solución.

TEOREMA 6. *Todo conjunto linealmente independiente de elementos de un espacio vectorial de dimensión finita, es parte de una base.*

Sea el conjunto independiente $\xi_1, \ldots, \xi_r$, y sea $\alpha_1, \ldots, \alpha_n$ una base de V. Formemos el conjunto $[\xi_1, \ldots, \xi_r, \alpha_1, \ldots, \alpha_n]$. Podemos extraer de aquí (Teorema 3, Corolario 2) un subconjunto independiente que también engendra V (o sea, una base para V), separando uno a uno los términos que sean combinación lineal de los que le preceden. Como los ξ_i son independientes, ningún ξ_i será separado y la base que resulte contendrá todos los ξ_i, c. q. d.

EJERCICIOS

1. Sean $\alpha_1, \alpha_2, \alpha_3$ tres vectores linealmente independientes en $V_3(F)$ tales, que cada vector unidad ϵ_i es una combinación lineal de α_1, α_2 y α_3. Demostrar que $\alpha_1, \alpha_2, \alpha_3$ forman una base de $V_3(F)$.
2. En Ejerc. 1 de § 4 ¿cuáles de los indicados conjuntos de vectores son bases del espacio que les contiene?
3. ¿Forman los vectores (1, 1, 0) y (0, 1, 1,) una base de $V_3(R)$?
4. En $V_4(R)$ hallar las coordenadas de los vectores unidad $\epsilon_1, \epsilon_2, \epsilon_3$ y ϵ_4, con relación a la base

   $\alpha_1=(1, 1, 0, 0)$, $\alpha_2=(0, 0, 1, 1)$, $\alpha_3=(1, 0, 0, 4)$, $\alpha_4=(0, 0, 0, 2)$.
5. Hallar las coordenadas de (1, 0, 1) relativas a la siguiente base en $V_3(C)$: $(2i, 1, 0)$, $(2, -i, 1)$, $(0, 1+i, 1-i)$.
6. Hallar cuatro vectores de $V_3(C)$ que en conjunto engendren un subespacio de dos dimensiones, siendo linealmente independientes dos a dos.
7. Mostrar que los números $a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}+e\sqrt{12}$, con $a, \ldots, e$ racionales, forman un anillo conmutativo, y que este anillo es un espacio vectorial sobre el campo racional R. Hallar una base para este espacio.
8. Mostrar geométricamente que dos vectores no colineales en $V_2(R^*)$ forman una base de $V_2(R^*)$.
9. Demostrar que:
   a) Un vector determinado de $V_2(F)$ nunca es una base de $V_2(F)$;
   b) Tres vectores de $V_2(F)$ son siempre linealmente dependientes (confróntese Ejerc. 2 de § 4).
   c) Cualquier base de $V_2(F)$ consiste, exactamente, en dos vectores.

10. En $V_4(R)$ hallar:

a) Una base que contenga al vector (1, 2, 1, 1);
b) Una base que contenga los vectores (1, 1, 0, 2) y (1, −1, 2, 0);
c) Una base que contenga los vectores (1, 1, 0, 0), (0, 0, 2, 2,) y (0, 2, 3, 0).

11. Si $c_1\alpha+c_2\beta+c_3\gamma=O$, con c_1 y $c_3\neq 0$, mostrar que α y β engendran el mismo subespacio que β y γ.

6. Dimensión

Los párrafos anteriores sugieren la idea de que cualquier ba de $V_n(F)$ tiene n elementos. Vamos a probar ahora esta «invaria cia del número dimensional» : demostraremos que todas las bases de un espacio vectorial V de dimensión finita tienen el mismo número de elementos.

En efecto, sea $A_0=[\alpha_1, \dots, \alpha_n]$ un conjunto de vectores que engendran V, y sea $[\xi_1, \dots, \xi_r]$ un subconjunto de V linealmente independiente. En tal caso, ξ_r es una combinación lineal de las α_i; y entonces, el conjunto $B_1=[\xi_r, \alpha_1, \dots, \alpha_n]$ engendra V y es linealmente dependiente. De aquí, por el Teorema 3, B_1 contendrá un término dependiente de sus predecesores. Si se prescinde de este término quedará un subconjunto $A_1=[\xi_r, \alpha_{i(1)}, \dots, \alpha_{i(n-1)}]$, que engendra V. Repitiendo el razonamiento, el conjunto $B_2=[\xi_{r-1}, \xi_r, \alpha_{i(1)}, \dots, \alpha_{i(n-1)}]$ engendra V y es linealmente dependiente. Hay, pues, algún término de B_2 que es combinación lineal de los precedentes. Como las ξ_i son independientes, este término no puede ser uno de ellos y deberá ser uno de los α_i. Si se suprime, nos encontraremos con otra sucesión de n elementos que engendra V, $A_2=[\xi_{r-1}, \xi_r, \alpha_{j(1)}, \dots, \alpha_{j(n-2)}]$. Repitiendo el argumento r veces llegaremos a una sucesión de n elementos que engendra V de la forma $A_r=[\xi_1, \dots, \xi_r, \alpha_{k(1)}, \dots, \alpha_{k(n-r)}]$. De esto concluímos que $n \geqslant r$. Este resultado podemos enunciarlo como sigue :

TEOREMA 7. *El número de elementos de cualquier conjunto que engendra un espacio vectorial* V *es igual o mayor que el de cualquier subconjunto linealmente independiente de* V.

En particular, si V tiene dos bases (dos conjuntos *independientes* que puedan engendrar V) constituídas por m y m' elementos respectivamente, se verificará a la vez que $m \geqslant m'$ y que $m' \geqslant m$. Tendrá, pues, que ser $m=m'$. Luego

COROLARIO 1. *Dos bases cualesquiera de un espacio vectorial de dimensión finita* V *tienen el mismo número de elementos.*

Este número, indicado por $d[V]$ se llama dimensión de V.

COROLARIO 2. *En un espacio vectorial* n*-dimensional,* $n+1$ *vectores cualesquiera son linealmente dependientes.*

COROLARIO 3. *Un espacio vectorial sobre* F *de dimensión finita es isomorfo con* $V_n(F)$ *para un solo y preciso valor de* n.

TEOREMA 8. *Para que un conjunto* A *de* n *vectores de* $V=V_n(F)$ *sea una base, es suficiente que puedan engendrar* V *o que sean linealmente independientes.*

Demostración. Si A engendra V, contiene un subconjunto A' que es una base (Teorema 3, Corolario 2) ; pero A' tiene n elementos por el Teorema 7, luego será igual a A. Asimismo, si A es independiente, será parte de una base A', por el Teorema 6, la cual tendrá n elementos, por el Teorema 7, y por tanto, será la misma A.

TEOREMA 9. *Si* S *y* T *son dos subespacios del espacio vectorial* V, *sus dimensiones satisfacen a*

$$d[S]+d[T]=d[S\frown T]+d[S+T]. \tag{23}$$

En efecto : sea $\xi_1, \ldots, \xi_n$ una base para $S\frown T$; por el Teorema 6, S y T tendrán bases $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_r$ y $\xi_1, \ldots, \xi_n, \zeta_1, \ldots, \zeta_s$, respectivamente. Evidentemente, las ξ_i, η_j y ζ_k juntas engendran $S+T$. Ellas constituirán una base, ya que

$$a_1\xi_1+\ldots+a_n\xi_n+b_1\eta_1+\ldots+b_r\eta_r+c_1\zeta_1+\ldots+c_s\zeta_s=O$$

implica que $\Sigma b_j\eta_j=-\Sigma a_i\xi_i-\Sigma c_k\zeta_k$ esté en T, de donde $\Sigma b_j\eta_j$ está en $S\frown T$ y entonces $\Sigma b_j\eta_j=\Sigma d_i\xi_i$ para algunos escalares d_i. Pero como los ξ_i y η_j son independientes, cada b_j es 0. Del mismo modo veríamos que cada $c_k=0$; por sustitución obtenemos $\Sigma a_i\xi_i=O$ y cada $a_i=0$. Esto nos muestra que los ξ_i, η_j y ζ_k son una base para $S+T$.

Después de demostrar esta propiedad, observemos que la conclusión es la de la regla de aritmética : $(n+r)+(n+s)=n+(n+r+s)$.

EJERCICIOS

1. Si dos subespacios S y T de un espacio vectorial V tienen la misma dimensión, demostrar que $S\leq T$ implica $S=T$.

2. En $V_4(R)$, dos subespacios S y T son engendrados, respectivamente, por los vectores

   S: (1, −1, 2, −3), (1, 1, 2, 0), (3, −1, 6, −6);
   T: (0, −2, 0, −3), (1, 0, 1, 0).

   Hallar las dimensiones de S, de T, de $S \frown T$ y de $S+T$.
3. Hallar la dimensión máxima posible de $S+T$ y la menor dimensión posible de $S \frown T$, donde S y T son subespacios variables de dimensiones determinadas s y t en $V_n(F)$.
*4. Demostrar para los subespacios que $S \frown T = S \frown T'$, $S+T=S+T'$ y $T \leqslant T'$ implican $T=T'$.
5. Un «automorfismo» de un espacio vectorial V significa un isomorfismo de V consigo mismo.
   a) Demostrar que la correspondencia $(x_1, x_2, x_3) \leftrightarrow (x_2, -x_1, x_3)$ es un automorfismo de $V_3(F)$.
   b) Demostrar que $V_n(F)$ tiene un automorfismo de «orden» n, en el sentido de la teoría de grupos.
6. Si S es un subespacio r-dimensional de $V_n(F)$, demostrar que existe un subespacio $(n-r)$-dimensional T tal, que $S \frown T = O$, $S+T=V_n(F)$.
7. Sea V un espacio vectorial con subespacios «complementarios» S y T satisfaciendo a $S \frown T = O$, $S+T=V$. Demostrar que V está determinado, salvo isomorfismos, por S y T.
8. Demostrar que, en Ejercicio 7, si V tiene una base finita, T estará determinada por S y V, salvo isomorfismos.
9. Demostrar que, dados los espacios vectoriales S y T sobre un campo F, existe un espacio vectorial V que tiene subespacios «complementarios» (Ejercicio 7) isomorfos con S y T, respectivamente.
10. Un automorfismo de $V_2(R)$ transforma (1, 0) en (0, 1) y (0, 1) en (−1, −1). ¿Cuál es su orden? ¿Depende la respuesta de la base del campo?
*11. Resolver Ejerc. 2 para el más general $V_4(J_p)$.
*12. Definir la «independencia lineal» de un conjunto infinito de vectores. Establecer algunos teoremas relativos a estos conjuntos independientes.
*13. a) ¿Cuántos conjuntos de dos elementos linealmente independientes tiene $V_3(J_2)$? ¿Cuántos de tres elementos? ¿Y de cuatro elementos?
   b) Generalizar la fórmula a $V_n(J_2)$ y a $V_3(J_p)$.
*14. Establecer una correspondencia biunívoca entre los automorfismos (confróntese Ejerc. 5) de un espacio vectorial y sus bases ordenadas. ¿Cuántos automorfismos tiene $V_n(J_p)$? Comprobarlo en el caso $n=p=2$.
*15. ¿Cuántos espacios k-dimensionales diferentes tiene $V_n(J_p)$?

7. Productos internos

La discusión precedente no ha hecho ninguna mención de la *longitud* de los vectores o de los *ángulos* que forman en el espacio ordinario. Para esto hay una razón poderosa.

Ciertamente, la correspondencia $\xi \to 2\xi$ dobla la longitud de cada vector en el plano, pero es un isomorfismo de $V_2(R^*)$ con res-

pecto a la adición vectorial y a la multiplicación por escalares. Esto nos dice (*) que no puede definirse la longitud mediante las operaciones vectoriales consideradas hasta ahora. Una observación análoga se aplica a la definición de ángulos.

Se puede, sin embargo, definir ambos conceptos de un modo conveniente mediante los productos internos. Llamaremos «producto interno» de dos vectores $\xi=(x_1, \ldots, x_n)$ y $\eta=(y_1, \ldots, y_n)$ con componentes reales, la cantidad

$$(24) \qquad (\xi, \eta)=x_1y_1+x_2y_2+\ldots+x_ny_n.$$

(Como esto es un escalar, los físicos llaman a este producto interno el «producto escalar» de dos vectores.) Los productos internos tienen cuatro propiedades algebraicas importantes, que son consecuencias inmediatas de la definición (24):

$$(25) \qquad (\xi+\eta, \zeta)=(\xi, \zeta)+(\eta, \zeta), \qquad (c\xi, \eta)=c(\xi, \eta);$$

$$(26) \qquad (\xi, \eta)=(\eta, \xi), \qquad (\xi, \xi)>0 \text{ a menos que } \xi=O.$$

Las dos primeras igualdades dicen que los productos internos son lineales con respecto al primer factor; la tercera expresa la ley simétrica o conmutativa. Del conjunto de estas tres resulta que el producto interno es lineal con respecto a ambos factores (bilinealidad); la cuarta condición es que el producto (ξ, ξ) sea positivo (excepto si $\xi=O$).

Así pues, la fórmula cartesiana que da la *longitud* (llamada también «valor absoluto» o «norma») $|\xi|$ de un vector, en el plano $V_2(R^*)$ es, precisamente, la raíz cuadrada de un producto interno:

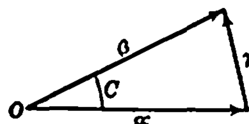


Figura 2

$$(27) \qquad |\xi|=(x_1^2+x_2^2)^{1/2}=(\xi, \xi)^{1/2}.$$

Una fórmula análoga se utiliza para calcular la longitud de un vector en el espacio tridimensional. Ahora bien, si α y β son dos vectores cualesquiera, el teorema trigonométrico del coseno aplicado al triángulo de lados α, β, $\gamma=\beta-\alpha$ (fig. 2) nos dará

$$|\beta-\alpha|^2=|\alpha|^2+|\beta|^2-2|\alpha|\cdot|\beta|\cdot\cos C$$

(*) Un razonamiento de este tipo se llama razonamiento «metamatemático». Ver, para más detalles, Cap. XI, § 7.

[con ángulo $C=\angle(\alpha, \beta)$]. Pero, por (25) y (27),

$$|\beta-\alpha|^2=(\beta-\alpha, \beta-\alpha)=(\beta, \beta)-2(\alpha, \beta)+(\alpha, \alpha).$$

Y teniendo en cuenta la anterior, resulta :

$$\cos \angle(\alpha, \beta)=(\alpha, \beta)/|\alpha|\cdot|\beta|. \tag{28}$$

El enunciado de esta igualdad es que el coseno del ángulo de dos vectores es igual al cociente de su producto interno por el producto de sus longitudes. De aquí se deduce que la condición para que dos vectores sean geométricamente *ortogonales* (o «perpendiculares») es que su producto interno sea nulo.

En vista de la facilidad con que los conceptos de adición de vectores y multiplicación por un escalar se pudieran generalizar a espacios de dimensión arbitraria sobre un campo arbitrario, es natural que tratemos ahora de generalizar de modo semejante los conceptos de longitud y ángulo. Sin embargo, nos encontramos con que, aunque aumentar el número de dimensiones no ofrece dificultad, aparecen inconvenientes al considerar ciertos campos. Los productos internos pueden definirse por (24), pero las longitudes

$$|\xi|=(\xi, \xi)^{1/2}=(x_1^2+x_2^2+\ldots+x_n^2)^{1/2} \tag{29}$$

no son definibles, a menos que toda suma de n cuadrados tenga una raíz cuadrada. Por otra parte, al querer conservar las propiedades de la distancia se tiene igualmente alguna dificultad.

Por estas razones, nos limitaremos en lo referente a longitudes, ángulos y distancias a considerar espacios vectoriales sobre el campo *real*. En el Cap. XI, § 12, trataremos las correspondientes nociones sobre el campo *complejo*.

EJERCICIOS

1. Demostrar en el plano, por geometría analítica, que el cuadrado de la distancia entre $\xi=(x_1, x_2)$ y $\eta=(y_1, y_2)$ es $|\xi|^2+|\eta|^2-2(\xi, \eta)$.
2. Utilizando los cosenos directores en el espacio tridimensional, demostrar que dos vectores ξ y η son ortogonales si, y sólo si, $(\xi, \eta)=0$.
3. Si la longitud se define por la fórmula (29) para los vectores ξ cuyos componentes sean números complejos, mostrar que existen vectores no nulos con longitud cero.
4. Mostrar que existe una suma de dos cuadrados que no tiene raíz cuadrada en el campo J_3 y en el J_7.

5. Demostrar las fórmulas (25) y (26), partiendo de la definición (24).
6. Demostrar la fórmula análoga a (25), estableciendo que el producto interno es lineal para el factor a la derecha.

8. Espacios vectoriales euclídeos abstractos

Nuestros razonamientos de geometría con cualquier número de dimensiones, se fundamentarán sobre la siguiente definición abstracta, sugerida por las consideraciones del § 7.

DEFINICIÓN. *Un espacio vectorial euclídeo es un espacio vectorial E con escalares reales, en el cual, a dos cualesquiera de sus vectores* ξ *y* η *corresponde un «producto interno» (real) que es simétrico, bilineal y positivo, en el sentido de (25) y (26).*

EJEMPLO 1. Todo $V_n(R^*)$ sobre el campo real es un espacio vectorial euclídeo n-dimensional si (ξ, η) se define por la igualdad (24).

EJEMPLO 2. Las funciones continuas reales $\phi(x)$ sobre el dominio $0 \leqslant x \leqslant 1$ forman un espacio vectorial euclídeo de número infinito de dimensiones, si definimos $(\phi, \psi) = \int_0^1 \phi(x)\psi(x)dx$

La *longitud* $|\xi|$ de un vector ξ, de un espacio vectorial euclídeo E, puede definirse mediante el producto interno, como raíz cuadrada positiva de (ξ, ξ); la existencia de esta raíz está asegurada por ser (ξ, ξ) positivo.

TEOREMA 10. *En un espacio vectorial euclídeo, la longitud tiene las siguientes propiedades:*

1) $|c\xi| = |c| \cdot |\xi|$;
2) $|\xi| > 0$, *excepto si es* $\xi = 0$;
3) $|(\xi, \eta)| \leqslant |\xi| \cdot |\eta|$ *(Desigualdad de Schwarz)*;
4) $|\xi + \eta| \leqslant |\xi| + |\eta|$ *(Desigualdad triangular).*

Demostraciones. Como $(c\xi, c\xi) = c^2(\xi, \xi)$, es cierta 1). La propiedad 2) es consecuencia de la condición de positividad exigida en la definición de un espacio vectorial euclídeo. La demostración de 3) es menos inmediata. En primer lugar utilizaremos el desarrollo

$$(\xi - \eta, \xi - \eta) = (\xi - \eta, \xi) - (\xi - \eta, \eta) =$$
$$= (\xi, \xi) - (\eta, \xi) - (\xi, \eta) + (\eta, \eta) =$$
$$= (\xi, \xi) - 2(\xi, \eta) + (\eta, \eta).$$

Por las condiciones de la definición, es $(\xi-\eta, \xi-\eta)\geqslant 0$; así, transponiendo, encontramos:

$$(\xi, \xi)+(\eta, \eta)=|\xi|^2+|\eta|^2\geqslant 2(\xi, \eta).$$

El mismo resultado obtendríamos reemplazando ξ por $-\xi$, de modo que $-2(\xi, \eta)$ es también menor o igual que $|\xi|^2+|\eta|^2$. Esto significa que el valor absoluto ordinario del número real $2(\xi, \eta)$ está acotado por

$$2(\xi, \eta)\leqslant|\xi|^2+|\eta|^2. \tag{30}$$

Si ξ' y η' son vectores de longitud 1, entonces $|\xi'|^2=|\xi'|\cdot|\eta'|=|\eta'|^2$, así que (30) da $2|(\xi', \eta')|\leqslant 2|\xi'|\cdot|\eta'|$, como se requiere en 3). También es trivial 3) en el caso de ser ξ o η igual a O.

Consideremos ahora vectores no nulos ξ y η. Por 1), los múltiplos escalares $\xi'=(1/|\xi|)\xi$ y $\eta'=(1/|\eta|)\eta$ tienen módulo 1, y en tal caso 3) es cierta para ξ' y η'. Pero ξ y η son los múltiplos escalares $\xi=c\xi'$, $\eta=d\eta'$ en los que $c=|\xi|$, $d=|\eta|$, así que

$$(\xi, \eta)=(c\xi', d\eta')=cd(\xi', \eta'),$$
$$|\xi|\cdot|\eta|=|c\xi'|\cdot|d\eta'|=|c|\cdot|d|(|\xi'|\cdot|\eta'|).$$

Esto nos muestra que los dos miembros de la deseada desigualdad 3) quedan multiplicados por el mismo factor $|cd|$ cuando ξ' y η' se multiplican por los escalares c y d. De lo cual deducimos que 3) es válida en general.

De la 3) obtenemos 4) fácilmente, pues

$$|\xi+\eta|^2=(\xi+\eta, \xi+\eta)=(\xi, \xi)+2(\xi, \eta)+(\eta, \eta)\leqslant$$
$$\leqslant|\xi|^2+2|\xi|\cdot|\eta|+|\eta|^2=(|\xi|+|\eta|)^2.$$

Si ahora decimos que la *distancia* entre dos elementos cualesquiera ξ y η de E es, *por definición*, el valor $|\xi-\eta|$, podemos demostrar que se cumplen las llamadas propiedades «métricas» de la distancia ordinaria, cuya consideración abstracta fué hecha por Frechet (1906).

TEOREMA 11. *La distancia tiene estas propiedades:* (M1), *la distancia de cada punto a sí mismo es nula, y entre dos puntos distintos es positiva;* (M2), *la distancia es simétrica, o sea,* $|\xi-\eta|=|\eta-\xi|$; (M3), *la distancia tiene la propiedad triangular* $|\xi-\eta|+|\eta-\zeta|\geqslant|\xi-\zeta|$.

Demostración. En primer lugar, $|\xi-\xi|=|O|=|0\cdot\xi|=0\cdot|\xi|=0$, por 1), mientras $|\xi-\eta|>0$ si $\xi-\eta\neq O$ (o $\xi\neq\eta$), por 2), lo cual prueba (*M*1). En segundo lugar, $|\xi-\eta|=|(-1)\cdot(\eta-\xi)|=|-1|\cdot|\eta-\xi|=|\eta-\xi|$, por 1), lo cual prueba (*M*2). Finalmente, (*M*3) se deduce de 4), pues

$$|\xi-\eta|+|\eta-\zeta|\geqslant|(\xi-\eta)+(\eta-\zeta)|=|\xi-\zeta|.$$

De la desigualdad de Schwarz deducimos en particular que, siendo ξ, η distintos de O, tenemos : $-1\leqslant(\xi,\eta)/|\xi|\cdot|\eta|\leqslant 1$. De aquí que $(\xi,\eta)/|\xi|\cdot|\eta|$ es el coseno de un solo ángulo comprendido entre 0° y 180°, al que *definimos* como *ángulo* de los dos vetores ξ y η [compárese con el caso particular (28)]. No demostraremos que los ángulos abstractos así definidos tengan las propiedades conocidas (¿podría probar el lector que $\angle(\xi,\eta)+\angle(\eta,\zeta)\geqslant\angle(\xi,\zeta)$?), excepto en el caso de ser ángulos *rectos*.

Dos elementos ξ y η se llaman *ortogonales* (en símbolos, $\xi\perp\eta$) cuando se verifica $(\xi,\eta)=0$. Esta definición, aplicada al anterior Ejemplo 2, nos proporciona un concepto analítico de gran importancia, que es el de «funciones ortogonales». Es fácil probar que si $\xi\perp\eta$, también $\eta\perp\xi$ (la relación de ortogonalidad es «simétrica»), y $c\xi\perp c'\eta$ para todos los c y c' (reales). También se observa que O es el único vector ortogonal a sí mismo. Además, si $(\eta,\xi_1)=\ldots=(\eta,\xi_m)=0$, se verifica que, para escalares cualesquiera c_i,

$$(\eta,\ c_1\xi_1+\ldots+c_m\xi_m)=c_1(\eta,\ \xi_1)+\ldots+c_m(\eta,\ \xi_m)=c_1\cdot 0+\ldots+c_m\cdot 0=0,$$

así que η es también ortogonal a cualquier combinación lineal de las ξ_i. Esto prueba :

TEOREMA 12. *Si un vector es ortogonal a $\xi_1,\ldots,\xi_m$, es también ortogonal a todos los vectores del subespacio engendrado por $\xi_1,\ldots,\xi_m$.*

EJERCICIOS

1. Sea $\xi=(1,2,3,4)$, $\eta=(0,3,-2,1)$. Calcular (ξ,η), $|\xi|$, $|\eta|$, $\angle(\xi,\eta)$.
2. Para ξ y η como en Ejerc. 1, hallar un vector de la forma $(1,1,0,0)+c_1\xi+c_2\eta$ ortogonal a ambos ξ y η.
3. a) ¿Son «ortogonales» sen $2\pi x$ y cos $2\pi x$, según el ejemplo 2 del texto?
   b) ¿Son ortogonales sen $2m\pi x$ y sen $2n\pi x$?

4. a) Hallar un polinomio de grado 2 «ortogonal», en el sentido dicho, a 1 y a x.
   b) Demostrar que existe un polinomio de grado n «ortogonal» a 1, x, x^2, ..., x^{n-1}.
5. Demostrar que en $V_3(R^*)$ hay, precisamente, dos vectores de longitud unidad, perpendiculares a dos vectores dados y linealmente independientes.
6. Demostrar que existe un vector con coordenadas racionales en $V_3(R^*)$ perpendicular a dos vectores dados cualquiera con coordenadas racionales.
7. Si α y $\beta \neq O$ son vectores fijos de un espacio vectorial euclídeo, hallar el vector más corto de la forma $\gamma = \alpha + t\beta$. ¿Es ortogonal a β? Dibujar una figura.
*8. Si α equidista de β y γ, demostrar que el punto medio del segmento $\overline{\beta\gamma}$ es el pie de la perpendicular desde α a $\overline{\beta\gamma}$.
*9. Probar que si $|\xi| = |\alpha|$ en un espacio vectorial euclídeo, entonces $\xi - \alpha \perp \xi + \alpha$; interpretarlo geométricamente.
*10. Sea un triángulo de vértices O, α, β. Calcular mediante la operación vectorial de producto interno los pies de las alturas del triángulo. Mostrar que estas alturas tienen un punto común.

•

9. Bases ortogonales y normales

En el Ejemplo 1 del § 8, los «vectores unitarios» $\epsilon_1 = (1, 0, ..., 0)$, ..., $\epsilon_m = (0, ..., 0, 1)$, tienen longitud igual a la unidad y son mutuamente ortogonales. Tenemos así un ejemplo de lo que llamaremos «base ortonormal».

DEFINICIÓN. *Los vectores* $\alpha_1, ..., \alpha_m$ *se llaman ortogonales y normales (o, brevemente, ortonormales) cuando: 1)* $|\alpha_i| = 1$ *para todo* i *; 2)* $\alpha_i \perp \alpha_j$ *si* $i \neq j$.

LEMA 1. *Los vectores* $\alpha_1, ..., \alpha_m$ *ortogonales y no nulos de un espacio vectorial euclídeo son linealmente independientes.*

Demostración. Si $x_1\alpha_1 + x_2\alpha_2 + ... + x_m\alpha_m = O$, para $k = 1, 2, ..., m$ se tendrá:

$$0 = (O, \alpha_k) = x_1(\alpha_1, \alpha_k) + ... + x_m(\alpha_m, \alpha_k) = x_k(\alpha_k, \alpha_k),$$

por la supuesta ortogonalidad. Pero $\alpha_k \neq O$ por hipótesis, luego $(\alpha_k, \alpha_k) > 0$ y $x_k = 0$.

Como corolario resulta que un conjunto de vectores ortonormales que engendren E, constituyen una base para E. Estas son las llamadas «bases ortonormales» de E.

LEMA 2. *Cualquier conjunto $\alpha_1, \ldots, \alpha_m$ de vectores ortonormales de un espacio vectorial euclídeo de dimensión finita, es parte de una base ortonormal.*

Demostración. Los α_i forman una base ortonormal para el subespacio S de E que ellos engendran. Si $S=E$, la conclusión es inmediata. En caso contrario, podemos encontrar un vector β de E que no esté en S. Procedamos a descomponer β en una parte «paralela» a S y una parte β^* perpendicular a S. A este fin pongamos:

(31) $$\beta^* = \beta - c_1\alpha_1 - \ldots - c_m\alpha_m, \text{ donde } c_k = (\alpha_k, \beta).$$

Entonces, con independencia de i, la propiedad lineal (25) da:

$$(\alpha_i, \beta^*) = (\alpha_i, \beta) - \sum_{k=1}^{m} c_k(\alpha_i, \alpha_k), \quad i=1, \ldots, m.$$

Pero $(\alpha_i, \beta)=c_i$, mientras (α_i, α_k) es cero si $i \neq k$ y 1 si $i=k$; de aquí que $(\alpha_i, \beta^*)=c_i-c_i=0$, lo que quiere decir que β^* es ortogonal a cada α_i. Como $\beta=\beta^*+\Sigma c_k\alpha_k$ no está en S, el conjunto $[\alpha_1, \ldots, \alpha_m, \beta]$ engendra un subespacio de E que contiene al S. Por la misma razón $\beta^* \neq O$, y también $|\beta^*|>0$: deducimos que el vector $\alpha_{m+1}=\beta^*/|\beta^*|$ tiene como longitud 1, y por ser un múltiplo escalar de β^* es ortogonal a los $\alpha_1, \ldots, \alpha_m$. Vemos que el conjunto $\alpha_1, \ldots, \alpha_m, \alpha_{m+1}$ es una base ortonormal para un subespacio T de dimensión $d[T]=m+1=d[S]+1$. La demostración se completa por inducción sobre el entero $d[E]-d[S]$.

Como siempre podemos encontrar un vector de longitud 1 en E, un corolario del Lema 2 es:

TEOREMA 13. *Todo espacio vectorial euclídeo E de dimensión finita tiene una base ortonormal.*

En particular, deducimos que todo subespacio m-dimensional S de un espacio vectorial euclídeo E tiene una base ortonormal $\alpha_1, \ldots, \alpha_m$. Si β es un vector cualquiera de E que no pertenezca al subespacio S, la construcción (31) descompone a β en la forma $\beta=\beta^*+\gamma$, en la que $\gamma=c_1\alpha_1+\ldots+c_m\alpha_m$ está en S, mientras β^* es ortogonal a cada vector α_i de la base, y por lo tanto (Teorema 12) a todos los vectores de S. Esta condición determina unívocamente β^* y γ, pues si en otra descomposición fuese $\beta'=\beta-\Sigma c_i'\alpha_i$ ortogonal a cada α_k, se tendría:

$$0=(\beta-\Sigma c_i'\alpha_i, \alpha_k)=(\beta, \alpha_k)-\sum_i c_i'(\alpha_i, \alpha_k)=(\beta, \alpha_k)-c_k'.$$

Los coeficientes $c_k' = (\beta, \alpha_k)$ son exactamente los mismos coeficientes c_k utilizados en (31). La conclusión es el siguiente

Corolario. *Si S es un subespacio de dimensión finita de un espacio vectorial euclídeo* E, *cualquier vector* β *de* E *puede expresarse de modo único como una suma* $\beta = \beta^* + \gamma$, *en la cual* γ *está en* S *y* β^* *es ortogonal a todos los vectores de* S.

El vector γ se llama *proyección ortogonal* de β sobre S.

Representemos ahora por S' el conjunto de todos los vectores de E ortogonales a todos los vectores de S (esto generaliza la noción de plano perpendicular a una recta dada). Si β_1 y β_2 están en S', entonces cada vector α de S es ortogonal a β_1 y β_2 y, por el Teorema 12, a toda combinación lineal $c_1\beta_1 + c_2\beta_2$. Por ello, S' contiene a todas las combinaciones lineales de sus vectores, luego es a su vez un subespacio; como el único vector ortogonal a sí mismo es O, S y S' tienen únicamente este vector común y $S \frown S' = O$. Por el corolario anterior, todo vector de E tiene la forma $\beta^* + \gamma$, en la que β^* está en S' y γ en S. Esto quiere decir que $S + S' = E$. La dimensión de S' puede ser calculada por (23). Esto se resume en el

Teorema 14. *Sea* S *un subespacio cualquiera de un espacio vectorial euclídeo* E. *El conjunto* S' *de los vectores ortogonales a todos y cada uno de los vectores de* S *es un subespacio que satisface a las relaciones* $\mathrm{S} \frown \mathrm{S}' = \mathrm{O}$, $\mathrm{S} + \mathrm{S}' = \mathrm{E}$ *y* $\mathrm{d}[\mathrm{S}] + \mathrm{d}[\mathrm{S}'] = \mathrm{d}[\mathrm{E}]$.

El subespacio S' se llama *complemento ortogonal* de S.

Concluiremos este capítulo determinando todos los productos internos posibles en un espacio vectorial euclídeo V (real).

Fácilmente se ve que si $\alpha_1, \ldots, \alpha_n$ es una base cualquiera de V, dados dos vectores arbitrarios $\xi = x_1\alpha_1 + \ldots + x_n\alpha_n$ y $\eta = y_1\alpha_1 + \ldots + y_n\alpha_n$, tendremos:

$$(32) \qquad (\xi, \eta) = (\Sigma x_i\alpha_i, \Sigma y_k\alpha_k) = \sum_{i,k} x_i y_k (\alpha_i, \alpha_k).$$

por la bilinealidad del producto interno. Así pues, el producto interno de dos vectores cualesquiera está determinado por las n^2 constantes reales $(\alpha_i, \alpha_k) = a_{ik}$ como una cierta forma «bilineal» $\sum_{i,k} a_{ik} x_i y_k$ en las coordenadas x_i e y_k. Como $(\alpha_i, \alpha_k) = (\alpha_k, \alpha_i)$, esta forma se llama «simétrica».

Recíprocamente, cualquier forma bilineal y simétrica $\sum_{i,k} a_{ik}x_iy_k$ ($a_{ik}=a_{ki}$), en $V_n(F)$, satisface las tres primeras condiciones de (25) y (26). La cuarta condición es que la forma *cuadrática* $\sum_{i,k} a_{ik}x_ix_k$ sea «definida positiva», es decir, sea positiva para todos los valores reales de las x_i, excepto cuando toda $x_i=0$. No es fácil traducir esto en una condición a la que deban cumplir los coeficientes a_{ik}.

Si suponemos la base ortonormal, tenemos $(\alpha_i, \alpha_k)=0$, si $i \neq k$, y $(\alpha_i, \alpha_i)=1$; de aquí que la (32) se reduce a

$$(33) \qquad (\xi, \eta)=\sum_{i=1}^{n} x_iy_i=x_1y_1+\ldots+x_ny_n.$$

TEOREMA 15. *Un producto interno «abstracto», referido a una base ortonormal, adopta la forma concreta (33).*

Con esto hemos terminado el análisis de los espacios vectoriales euclídeos abstractos de número finito de dimensiones ; únicamente continúa en el misterio la naturaleza de los de infinitas dimensiones.

EJERCICIOS

1. Hallar bases ortogonales normales de los subespacios de un espacio euclídeo de cuatro dimensiones engendrados por
   a) (1, 1, 0, 0), (0, 1, 2, 0) y (0, 0, 3, 4);
   b) (2, 0, 0. 0), (1, 3, 3, 0) y (0, 4, 6, 1).
   (*Sugerencia: Primero* hallar bases ortogonales, *después* normalizarlas.)
2. Dibujar una figura que ilustre la proyección ortogonal de un vector sobre un subespacio unidimensional.
3. Hallar las proyecciones ortogonales de $\beta=(2, 1, 3)$ sobre el subespacio engendrado por $\alpha=(1, 0, 1)$.
4. Hallar la proyección ortogonal de $\beta=(0, 0, 0, 3)$ sobre cada uno de los subespacios del Ejerc 1.
5. Hallar el complemento ortogonal de los subespacios engendrados por (2, 1, —2) en un espacio euclídeo tridimensional.
6. Hallar el complemento ortogonal de cada uno de los subespacios del Ejercicio 1.
7. Demostrar que para dos subespacios cualesquiera S y $T \geqslant S$ de un espacio vectorial euclídeo E de finitas dimensiones, existe un subespacio X que es un «complemento relativo» de S en T, en el sentido de que a) $\xi \in X$ y $\eta \in S$ implica $\xi \perp \eta$, b) $S+X=T$. Demostrar que este subespacio es único, y discutir sus dimensiones.
8. Demostrar que el complemento ortogonal S' de un espacio dado S en E es un subespacio completamente determinado por la propiedad de que cualquier vector de S' es ortogonal a cualquier vector de S y que $S+S'=E$.

9. Demostrar que $(S')'=S$; en otras palabras, la relación «T es complemento ortogonal de S» es una relación simétrica entre subespacios.

*10. Demostrar $(S+T)'=S' \cap T'$ y $(S \cap T)'=S'+T'$ en cualquier espacio euclídeo de finitas dimensiones. (*Advertencia:* Una parte es más fácil de demostrar que la otra.)

11. Demostrar que una base ortogonal normal para S y una base ortogonal normal para S' constituyen, juntas, una base ortogonal normal para E.

CAPÍTULO VIII

Algebra de las matrices

1. Transformaciones lineales y matrices

Hay muchas y muy conocidas maneras de representar linealmente el plano sobre sí mismo, es decir, de hacer que cada punto (x, y) del plano se transforme en un nuevo punto, cuyas coordenadas (x', y') vengan dadas como funciones lineales y homogéneas de x e y :

(1) $$x' = xa + ya^*, \qquad y' = xb + yb^*$$

Por ejemplo: la *rotación* con centro en el origen y de amplitud θ (medida en sentido contrario a las agujas de un reloj) se expresa por las ecuaciones lineales

(2) $$R_\theta : \quad x' = x \cos \theta - y \operatorname{sen} \theta, \quad y' = x \operatorname{sen} \theta + y \cos \theta.$$

Si el plano se dilata uniforme y radialmente desde el origen, moviéndose cada punto (x, y) hasta una distancia de dicho origen cuya razón con la distancia primitiva sea $k > 1$, las nuevas coordenadas del punto (x', y') son asimismo funciones lineales de las antiguas,

(3) $$S_k : \qquad x' = kx, \qquad y' = ky.$$

Una transformación de la forma (3) será llamada transformación de *semejanza*. Si $0 < k < 1$, se tratará de una *compresión* hacia el origen.

Si $k = -1$, se tiene una simetría respecto al origen, o una rotación de 180° ($x' = -x$, $y' = -y$). También es una transformación

lineal del plano, el movimiento de *cizalla o corrimiento sobre el eje* x, cuyas ecuaciones son:

(4) $$H_a: \quad x' = x + ay, \quad y' = y.$$

En esta transformación, cada punto se mueve sobre una paralela al eje x, y un rectángulo de lados paralelos a los ejes es transformado en un paralelogramo (figura 1). También mediante una transformación lineal, el plano se dilata de modo uniforme a lo largo del eje x,

(5) $$U_b: \quad x' = bx, \quad y' = y \quad (b > 1).$$

Si $0 < b < 1$, estas ecuaciones representan una compresión hacia el eje y. Si $b = -1$, se tendrá una *simetría* o *reflexión* respecto al eje y. Existen, además, transformaciones lineales de todo el plano sobre una parte del mismo. Una de ellas es la $x' = x$, $y' = 0$, que proyecta cada punto sobre el eje x. Evidentemente, esta transformación no es biunívoca.

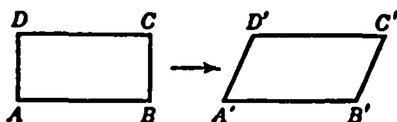


Figura 1

Los precedentes ejemplos se han interpretado como sendas transformaciones de los puntos P de coordenadas (x, y), pero evidentemente pueden interpretarse como transformaciones de los correspondientes vectores OP de componentes x e y. En general, las ecuaciones (1) definen una transformación T que transforma al vector $\xi = (x, y)$ en el vector $\xi' = (x', y') = (xa + ya^*, xb + yb^*)$, esto es: $\xi \to \xi T = \xi'$. Las transformaciones lineales tienen gran número de propiedades algebraicas extremadamente importantes.

Por ejemplo, si x e y se multiplican por un mismo escalar c, veremos que, por (1), x' e y' quedarán asimismo multiplicados por c, de modo que el vector $(c\xi)T$, transformado del producto de un escalar por un vector, es $c\xi' = c(\xi T)$, o sea, el producto del escalar por el transformado del vector. De un modo análogo, el vector transformado de la suma de dos vectores $(x_1 + x_2, y_1 + y_2)$, es la suma de los vectores transformados de los sumandos (x_1, y_1) y (x_2, y_2), según resulta inmediatamente de (1). Estas dos propiedades,

(6) $$(c\xi)T = c(\xi T), \quad (\xi_1 + \xi_2)T = \xi_1 T + \xi_2 T$$

son suficientes para dar una caracterización abstracta de las transformaciones lineales homogéneas (1), independientemente de la base del espacio vectorial.

En efecto, supongamos una transformación ξT que esté definida para todo vector $\xi=(x, y)$, de modo que se cumplan las condiciones (6). Sean $\epsilon_1=(1, 0)$ y $\epsilon_2=(0, 1)$ los vectores unitarios, y sean $\epsilon_1 T=(a, b)$ y $\epsilon_2 T=(a^*, b^*)$ sus respectivos transformados. Puesto que $(x, y)=x\epsilon_1+y\epsilon_2$, las condiciones (6) nos dan :

$$(x, y)T=(x\epsilon_1+y\epsilon_2)T=(x\epsilon_1)T+(y\epsilon_2)T=$$
$$=x(a, b)+y(a^*, b^*)=(xa+ya^*, xb+yb^*).$$

Las componentes $x'=xa+ya^*$, $y'=xb+yb^*$ del vector transformado son, pues, funciones lineales de x e y como en las ecuaciones (1). Por esto, las propiedades (6) implican que la transformación está representada por ecuaciones de la forma (1) y recíprocamente.

El precedente estudio de las transformaciones lineales del plano real puede generalizarse a espacios de cualquier número de dimensiones. Tampoco es preciso que las coordenadas sean números reales, sino que pueden ser elementos de cualquier campo F.

DEFINICIÓN. *Una transformación lineal* T *de un espacio vectorial* V *es una correspondencia que representa a todo vector* ξ *de* V *sobre un vector* ξT *de* V, *de tal modo, que para todos los vectores* ξ *y* η *y todos los escalares* c *y* d *es*

$$(\mathrm{c}\xi+\mathrm{d}\eta)\mathrm{T}=\mathrm{c}(\xi\mathrm{T})+\mathrm{d}(\eta\mathrm{T}). \tag{7}$$

Como en el caso del plano, esta definición implica que la transformación T pueda representarse por un conjunto de ecuaciones lineales homogéneas. En efecto, cada vector ϵ_i de una base dada $\epsilon_1, \ldots, \epsilon_n$, tiene un transformado $\epsilon_i T$, el cual, lo mismo que cualquier vector del espacio, puede ser expresado como una combinación lineal de los vectores de la base, o sea :

$$
\begin{aligned}
\epsilon_1 T &= a_{11}\epsilon_1 + a_{12}\epsilon_2 + \ldots + a_{1n}\epsilon_n = \sum_j a_{1j}\epsilon_j \\
\epsilon_2 T &= a_{21}\epsilon_1 + a_{22}\epsilon_2 + \ldots + a_{2n}\epsilon_n = \sum_j a_{2j}\epsilon_j \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\epsilon_n T &= a_{n1}\epsilon_1 + a_{n2}\epsilon_2 + \ldots + a_{nn}\epsilon_n = \sum_j a_{nj}\epsilon_j .
\end{aligned} \tag{8}
$$

Un vector arbitrario $\xi = \sum_i x_i \epsilon_i$ tendrá, pues, un transformado que podemos expresar, utilizando repetidamente la (7), como sigue:

$$(\sum_i x_i \epsilon_i) T = \sum_i x_i(\epsilon_i T) = \sum_{i,j} x_i a_{ij} \epsilon_j = \sum_j (\sum_i x_i a_{ij}) \epsilon_j.$$

En la última expresión hemos agrupado los coeficientes de cada vector ϵ_j en un solo término $\Sigma x_i a_{ij}$. Este coeficiente es, por lo tanto, con relación a la base $\epsilon_1, \ldots, \epsilon_n$, la coordenada y_j del vector transformado $\xi T = \eta = \Sigma y_j \epsilon_j$. En otras palabras,

$$
\begin{aligned}
y_1 &= x_1 a_{11} + x_2 a_{21} + \ldots + x_n a_{n1} = \sum_i x_i a_{i1} \\
y_2 &= x_1 a_{12} + x_2 a_{22} + \ldots + x_n a_{n2} = \sum_i x_i a_{i2} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
y_n &= x_1 a_{1n} + x_2 a_{2n} + \ldots + x_n a_{nn} = \sum_i x_i a_{in}
\end{aligned}
\tag{9}
$$

Estas ecuaciones *representan* la transformación T referida a una base dada, en el sentido de que T hace corresponder al vector ξ de coordenadas $(x_1, \ldots, x_n)$ el vector $\eta = \xi T$, cuyas coordenadas $(y_1, \ldots, y_n)$ están dadas por las ecuaciones (9).

TEOREMA 1. *Con referencia a una base dada $\epsilon_1, \ldots, \epsilon_n$, del espacio vectorial V, cualquier transformación lineal T de V sobre sí mismo es representable por un conjunto (9) de ecuaciones lineales homogéneas en las coordenadas.*

Los coeficientes a_{ij} que intervienen en las ecuaciones (8) y (9) pueden ser dispuestos según la ordenación que se muestra a la derecha. Estos a_{ij} pertenecerán al campo F, si F es el campo de los escalares del espacio vectorial dado. La disposición de estos $n \times n$ coeficientes en un cuadro A constituye una *matriz cuadrada* de *orden* n sobre F (*), o, como también se dice, una matriz $n \times n$ sobre F. La fila i-ésima de la matriz es:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

$$A_i = (a_{i1}, a_{i2}, \ldots, a_{in}). \tag{10}$$

(*) Si el lector desea considerar sólo el caso importante en el que estos escalares son los números reales, en todo el capítulo deberá leer «conjunto de todos los números reales», en vez de «campo F».

Esta es, simplemente, la i-ésima fila de (8) o, lo que es lo mismo, la i-ésima columna de (9). De modo análogo, los elementos a_{ij}, con j fijo, dan la *columna* j-ésima de A, designada por $A^{(j)}$. Esta es también la fila j de (9); de modo que las filas de (9) son las columnas de (8) (*).

La matriz completa puede ser representada abreviadamente por $A = \| a_{ij} \|$. Cada matriz está unívocamente determinada por las correspondientes ecuaciones lineales homogéneas (9); por lo tanto,

TEOREMA 2. *Con referencia a una base dada $\epsilon_1, \ldots, \epsilon_n$ de un espacio vectorial $V_n(F)$, existe una correspondencia biunívoca entre las transformaciones lineales T de V y las matrices $n \times n$ sobre F. La matriz $\| a_{ij} \|$ corresponde a la transformación (9) con coeficientes a_{ij} transpuestos.*

Por ejemplo, en el plano, la rotación, semejanza y homología dadas por (2), (3) y (4) corresponden respectivamente a las matrices

$$R_\theta \rightarrow \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}, \qquad S_k \rightarrow \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \qquad H_a \rightarrow \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}.$$

EJERCICIOS

1. Explicar la interpretación geométrica de cada una de las siguientes transformaciones lineales:

   a) $y' = x,\ x' = y$; b) $y' = x,\ x' = x$;
   c) $y' = x,\ x' = 0$; d) $y' = ky,\ x' = kx + kay$;
   e) $y' = by,\ x' = cx$.

2. Considerar la transformación que transforma cada punto P en otro P' relacionado con el P del modo descrito a continuación; determinar, además, cuándo la transformación es lineal, y hallar sus ecuaciones:
   a) P' está dos unidades a la derecha de P y una unidad encima (traslación).
   b) P' es la proyección de P sobre la recta de inclinación 1/2 que pasa por el origen.
   c) P' está en la semirrecta OP que une P con el origen, y a una distancia de O tal, que $OP' = 4/OP$.

(*) Nuestra decisión de ordenar los coeficientes a_{ij} en el orden en que aparecen en (8), con preferencia al orden «transpuesto» de (9), es puramente *convencional*. La hemos adoptado teniendo en cuenta la convención usual para el producto de matrices (§ 2) y nuestra costumbre de indicar la transformación ξT a la *derecha* del vector. Habríamos empleado la segunda ordenación si la transformación se escribiese a la izquierda, como en $T(\xi)$.

d) P' se obtiene del P por una rotación de 30° alrededor del origen, seguida de un movimiento del tipo (4), pero paralelo al eje y.
e) P' es el simétrico de P respecto a la recta $x=3$.

3. Hallar las ecuaciones de las siguientes transformaciones en el espacio de tres dimensiones. ¿Cuáles son lineales?
a) Una transformación de semejanza.
b) Un movimiento de cizalladura paralelo al eje x.
c) Una reflexión en el plano x, y.
d) Una traslación.
e) Una rotación alrededor del eje y.

4. Describir los efectos geométricos de las siguientes transformaciones lineales del espacio:
a) $x'=ax$, $y'=by$, $z'=cz$;
b) $x'=0$, $y'=3y$, $z'=3z$;
c) $x'=x+2y+5z$, $y'=y$, $z'=z$;
d) $x'=x-y$, $y'=x+y$, $z'=4z$.

5. Probar que la transformación lineal T de V_n está completamente determinada cuando se conocen los transformados $\alpha_1 T, \ldots, \alpha_n T$, de n vectores α_i linealmente independientes.

6. Hallar las matrices que representan las transformaciones lineales determinadas como sigue:
a) $(1, 1) \to (0, 1)$ y $(-1, 1) \to (3, 2)$;
b) $(1, 0) \to (4, 0)$ y $(0, 1) \to (-1, 2)$;
c) $(2, 3) \to (1, 0)$ y $(3, 2) \to (1, -1)$;
d) $(1, 0, 0) \to (1, 2, 1)$, $(0, 1, 0) \to (3, 1, 1)$, $(0, 0, 1) \to (0, 0, 3)$.

7. La *imagen* de un subespacio S de V debida a una transformación lineal T es el conjunto $(S)T$ de todos los vectores ξT, para ξ en S. Probar que $(S)T$ es también un subespacio.

2. Operaciones sobre matrices

El álgebra de las operaciones lineales (matrices) comprende tres operaciones: la adición de dos transformaciones lineales (o matrices), la multiplicación de una transformación lineal por un escalar y la multiplicación de dos transformaciones lineales. La combinación más importante de dos transformaciones lineales T y U es su producto TU (se aplica primero T y después U, como en la teoría de grupos). Formalmente, $\xi(TU)$ es, pues, una transformación definida como $(\xi T)U$ para cada vector ξ.

Por ejemplo, si el corrimiento H_a de (4) va seguido por una semejanza S_k que transforme (x', y') en $x''=kx'$, $y''=ky'$, el efecto combinado de ambas es transformar (x, y) en $x''=kx+kay$, $y''=ky$. El producto H_aS_k es asimismo lineal.

TEOREMA 3. *El producto de dos transformaciones lineales es también una transformación lineal.*

Demostración. Por definición, un producto TU transforma un vector ξ en un vector $\xi(TU)=(\xi T)U$. Por la linealidad de T y de U es, sucesivamente,

$$(11) \qquad (c\xi+d\eta)TU=[c(\xi T)+d(\eta T)]U=c(\xi TU)+d(\eta TU),$$

lo que nos dice que TU satisface a la condición (7), característica de las transformaciones lineales.

Lo anterior implica que las ecuaciones lineales y homogéneas que representan a T y U pueden ser combinadas para obtener las ecuaciones lineales y homogéneas que representen a TU. Para concretar, sea T la transformación

$$(12) \qquad \begin{aligned} x' &= xa_{11}+ya_{21}, \\ y' &= xa_{12}+ya_{22}, \end{aligned} \qquad A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

con matriz A, y apliquemos, después de ella, una transformación U, tal que (x', y') se transforme en (x'', y'') con

$$(13) \qquad \begin{aligned} x'' &= x'b_{11}+y'b_{21}, \\ y'' &= x'b_{12}+y'b_{22}, \end{aligned} \qquad B=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

La transformación producto, que resulta por sustitución de (12) en (13), es

$$(14) \qquad \begin{aligned} x'' &= (a_{11}b_{11}+a_{12}b_{21})x+(a_{21}b_{11}+a_{22}b_{21})y, \\ y'' &= (a_{11}b_{12}+a_{12}b_{22})x+(a_{21}b_{12}+a_{22}b_{22})y. \end{aligned}$$

La matriz de los coeficientes de la transformación producto se deduce de las matrices dadas A y B, por la importante regla

$$(15) \qquad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}.$$

(o dicho con brevedad : «multiplicando filas por columnas»). Así, el elemento de la *primera fila* y la *segunda columna* del resultado, depende sólo de la *primera fila* de A y de la *segunda columna* de B. Esta regla para multiplicar transformaciones lineales homogéneas resulta muy cómoda, pues evita el uso de las variables, que es necesario en las (14).

Para las matrices $n\times n$ se obtienen fórmulas análogas. Sea $\xi T=\eta$ una transformación, dada por las ecuaciones (9) relativamente a la base $\epsilon_1, \ldots, \epsilon_n$, y sea U una segunda transformación, que hace corresponder a η el vector $\zeta=\eta U$ según las ecuaciones

$$(16)\qquad z_k=y_1b_{1k}+\ldots+y_nb_{nk}=\sum_j y_jb_{jk},\qquad k=1,\ \ldots,\ n.$$

El producto TU transforma sucesivamente ξ en η y η en ζ. Las coordenadas z_k de ζ se hallarán por sustitución de (9) en (16), obteniendo

$$(17)\qquad \begin{aligned} z_k&=(\sum_i x_ia_{i1})b_{1k}+\ldots+(\sum_i x_ia_{in})b_{nk}=\sum_j(\sum_i x_ia_{ij})b_{jk}=\\ &=\sum_{i,j} x_ia_{ij}b_{jk}=x_1(\sum_j a_{1j}b_{jk})+\ldots+x_n(\sum_j a_{nj}b_{jk}).\end{aligned}$$

Así resulta probado otra vez que la transformación TU es lineal. Su representación analítica está dada por las ecuaciones $z_k=\sum_i x_ic_{ik}$, en donde los coeficientes c_{ik} son, según (17),

$$(18)\qquad c_{ik}=a_{i1}b_{1k}+a_{i2}b_{2k}+\ldots+a_{in}b_{nk}=\sum_{j=1}^{n} a_{ij}b_{jk},$$

$$(i=1,\ \ldots,\ n;\ k=1,\ \ldots,\ n).$$

La matriz $\| c_{ik}\|$ de la transformación producto, dada por (18), es llamada *matriz producto* de $A=\| a_{ij}\|$ por $B=\| b_{jk}\|$,

$$(19)\qquad \| c_{ik}\|=\| a_{ij}\|\cdot\| b_{jk}\|,\ \text{con}\ c_{ik}=\sum_{j=1}^{n} a_{ij}b_{jk}\ \text{para todo}\ i\ \text{y}\ k.$$

TEOREMA 4. *Con referencia a una base dada de* V, *las transformaciones lineales* T *y* U, *representadas respectivamente por matrices* A *y* B, *tienen un producto* TU, *representado por la matriz producto* AB.

El producto de dos matrices se efectúa, pues, así: el elemento c_{ik} de la fila i y columna k es el producto de la fila i de A (primer factor) por la columna k de B, entendiéndose que el «producto de una fila por una columna» es la suma de los productos de cada dos elementos homólogos.

La multiplicación de *matrices* es asociativa, ¡pero no es preciso probarlo desarrollando los cálculos a partir de la definición (19)! Cada matriz producto AB representa el producto de dos transfor-

maciones TU (Teor. 4). El producto de transformaciones es asociativo, como se vió en § 2 del Cap. VI. Por lo tanto, el producto de matrices es asociativo, lo que equivale a decir que $(AB)C = A(BC)$.

En correspondencia con la transformación idéntica, está la *matriz idéntica* $n \times n$, I, llamada también *matriz unidad*, la cual consta de sólo elementos 1 en la diagonal principal (que va del ángulo superior izquierda al inferior derecha), siendo ceros los restantes. Como representa la transformación idéntica, tiene la propiedad de que $IA = A = AI$, siendo A cualquier matriz $n \times n$. Por ejemplo.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} 1\cdot a_1 + 0\cdot a_2 & 1\cdot b_1 + 0\cdot b_2 \\ 0\cdot a_1 + 1\cdot a_2 & 0\cdot b_1 + 1\cdot b_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

La multiplicación no es conmutativa, y así

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(*Sugerencia:* ¿Qué transformaciones geométricas producen estas matrices sobre el cuadrado de Cap. VI, § 1?)

No todas las matrices tienen inversa respecto a la multiplicación. En los productos

(20)
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} 3a_2 & 3b_2 \\ 0 & 0 \end{pmatrix}$$

la segunda columna es siempre 0, así que el producto nunca puede ser I. Por lo tanto, toda matriz 2×2 cuya segunda fila consista en ceros, no puede tener inversa. Tampoco es válida la ley de simplificación, puesto que hay numerosos divisores de cero, como en

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Pasamos ahora a definir la adición de matrices. La *matriz suma* de dos matrices $n \times n$ se obtiene sumando los elementos homólogos de ambas, es decir,

(21)
$$\| a_{ij} \| + \| a_{ij}^* \| = \| a_{ij} + a_{ij}^* \|.$$

Esta suma satisface las leyes asociativa y conmutativa, porque los términos a_{ij} las satisfacen también. La matriz que tiene todos sus elementos nulos, se comporta en la adición como la *matriz cero*. La matriz inversa de otra A, respecto a la adición, resulta anteponiendo el signo menos a cada elemento de A. Las matrices forman, pues, un grupo abeliano respecto a la suma.

El *producto escalar* cA de una matriz A por un escalar c, se obtiene multiplicando por c cada uno de los elementos de A. Es fácil demostrar las leyes usuales,

$$1 \cdot A = A, \qquad (cd)A = c(dA), \qquad c(A + B) = cA + cB,$$
$$(c + d)A = cA + dA.$$

TEOREMA 5. *Con la introducción de la adición y de la multiplicación escalar, todas las matrices* n × n *con elementos de un campo* F, *forman un espacio vectorial sobre* F.

Cualquier matriz $\|a_{ij}\|$ puede ser escrita como una suma $\sum_{i,j} a_{ij}E_{ij}$, donde E_{ij} es una matriz especial, cuyos elementos son todos ceros excepto el de la columna i y fila j, que es un 1. Estas matrices son linealmente independientes, luego forman una base para el espacio de las matrices $n \times n$. La dimensión de este espacio es n^2.

Entre las transformaciones lineales T y U de un espacio V, hay las dos operaciones correspondientes. La *suma* $T + U$ es la transformación que hace corresponder a ξ el vector $\xi(T + U) = \xi T + \xi U$. El producto escalar cT es la transformación $\xi(cT) = c(\xi T)$. La suma $T + U$ es lineal según (7), pues

$$(c\xi + d\eta)(T + U) = (c\xi + d\eta)T + (c\xi + d\eta)U = c\xi(T + U) + d\eta(T + U).$$

El producto cT es también lineal. Las ecuaciones lineales homogéneas representantes de la suma $T + U$ se hallan sumando los segundos miembros de las ecuaciones de T y de U respectivamente, esto es, por la adición según (21) de las matrices correspondientes. Para una base dada, la correspondencia de estas operaciones es exacta : si T y U se corresponden con las matrices A y B, entonces $T + U \leftrightarrow A + B$, $cT \leftrightarrow cA$: el espacio vectorial de las transformaciones T es isomorfo con el espacio vectorial de las matrices A.

La multiplicación de matrices no sólo es asociativa : también

es *distributiva* respecto a la suma. Pues la matriz $(A+A^*)B$ tiene sus elementos dados por las fórmulas del tipo (18),

$$d_{ik}=\sum_j (a_{ij}+a^*_{ij})b_{jk}=\sum_j a_{ij}b_{jk}+\sum_j a^*_{ij}b_{jk}.$$

Esto quiere decir que d_{ik} es la suma de un elemento c_{ik} de AB y el homólogo c^*_{ik} en A^*B, luego prueba la primera de las dos leyes

(22) $$(A+A^*)B=AB+A^*B,\quad A(B+B^*)=AB+AB^*.$$

Para los productos por un escalar d, se pueden también comprobar las leyes

(23) $$(dA)B=d(AB),\quad A(dB)=d(AB).$$

Las leyes (22) y (23) se sintetizan diciendo que la multiplicación de matrices es *bilineal*, pues combinando las primeras igualdades de estas leyes, resulta que $(dA+d^*A^*)B=d(AB)+d^*(A^*B)$. Y esta es exactamente la condición para que la multiplicación por B a la derecha sea una transformación lineal $X\to XB$ sobre el espacio vectorial de todas las matrices $n\times n$. El resto de las leyes (22) y (23) aseguran que la multiplicación por A a la izquierda es también una transformación lineal.

Resumiendo:

TEOREMA 6. *El conjunto de todas las matrices* n × n *sobre un campo* F, *es cerrado para la multiplicación; ésta es asociativa, tiene un elemento idéntico y es bilineal con relación a la adición y a la multiplicación escalar.*

Apéndice. La noción de transformación lineal puede también aplicarse a espacios de infinitas dimensiones.

EJEMPLO 1. Sea V el espacio constituído por todas las funciones $f(x)$ de una variable real x y sea J la transformación u «operador» $[f(x)]J=f(x+1)$. Si I es la transformación idéntica, el operador $\Delta=J-I$ es llamado un operador de diferencias; hace corresponder a $f(x)$ la función $f(x+1)-f(x)$. Ambos, J y Δ, son lineales, pues $[cf(x)+dg(x)]J=c[f(x)]J+d[g(x)]J$. Esta definición de linealidad es la aplicada antes, pero obsérvese que en este espacio infinito no podemos formular las correspondientes ecuaciones lineales homogéneas. Para una $a(x)$ dada, la operación $f(x)\to a(x)f(x)$ es también lineal.

EJEMPLO 2. Sea D el operador derivada, aplicado al espacio $C^{(\infty)}$ de todas las funciones $f(x)$ que poseen derivadas de todos los órdenes, el cual transforma $f(x)$ en $f'(x)$; D es un operador lineal. El teorema de Taylor se puede escribir simbólicamente como $e^D=J$.

EJEMPLO 3. Para las funciones $f(x, y)$ de dos variables, existen los correspondientes operadores lineales J_x, J_y, D_x, D_y, Δ_x, Δ_y. Así, $[f(x, y)]J_x=f(x+1, y)$; $[f(x, y)]D_x=f_x'(x, y)$; etc.

EJERCICIOS

1. Hallar los productos indicados, para las matrices

$$A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix},\quad B=\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix},\quad C=\begin{pmatrix}1 & 3\\ 2 & 1\end{pmatrix},\quad D=\begin{pmatrix}4 & 0\\ 2 & 1\end{pmatrix};$$

a) AB, BA, $A^2+AB-2B$;
b) $(A+B-I)(A-B+I)-(A+2B)(B-A)$;
c) DB, AC, AD;
d) Comprobhr la ley asociativa para los productos $(AC)D$, $A(CD)$.

2. Empleando el producto de matrices, hallar las ecuaciones de las siguientes transformaciones (notaciones de (2)-(5) de §1).

a) U_bH_a; b) S_kU_b; c) $R\ U_b$, con $\theta=45°$;
d) $R\ H_aU_b$, con $\theta=30°$; e) $S_kH_aU_bS_k$.

3. ¿Cuándo es $H_aS_k=S_kH_a$? (Notación como en Ejerc. 2.)
4. En el Ejercicio 4 del § 1, designamos por T_n la transformación descrita en el apartado (n). Calcular (usando matrices) los siguientes productos:

a) T_bT_c; b) T_aT_c; c) $T_bT_aT_b$; d) T_dT_c; e) $T_cT_bT_d$.

5. Probar las leyes (23) y la segunda mitad de (22).
6. a) Desarrollar $(A+B)^3$; b) Probar que $A^3A^2=A^2A^3$.
7. Demostrar la ley asociativa del producto de matrices directamente, a partir de (19).
8. Considerar un nuevo «producto» $A\times B$ de dos matrices, definido por el producto de «filas por filas». ¿Es asociativo?
9. a) Calcular los productos BE_1, BE_2, BE_3, E_2E_3, E_1E_2, donde

$$B=\begin{pmatrix}1 & 2 & 1\\ 1 & 3 & 2\\ 1 & 4 & 6\end{pmatrix},\quad E_1=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix},\quad E_2=\begin{pmatrix}1 & 0 & k\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},\quad E_3=\begin{pmatrix}a & 0 & 0\\ 0 & b & 0\\ 0 & 0 & c\end{pmatrix}.$$

b) Si A es una matriz 3×3, ¿cómo es AE_1 respecto a A?
c) Describir los efectos causados por multiplicación de una matriz por E_2; por E_3.

En los ejercicios 10-13, entiéndase que una matriz *diagonal* $D=\|d_{ij}\|$ es una matriz que tiene todos sus elementos nulos, excepto los de la diagonal principal (esto es, que $d_{ij}=0$ para $i\neq j$).

10. Probar que el conjunto de todas las matrices diagonales es cerrado para la multiplicación. ¿Este conjunto es un dominio de integridad? ¿Es un anillo conmutativo con unidad?
11. ¿Cuál es el efecto de multiplicar una matriz $n \times n$, A, por una matriz diagonal D? [Comparar Ejerc. 9, b).]
12. Hallar todas las matrices permutables con la matriz E, del Ejerc. 9. (Suponemos que a, b y c son distintos.)
13. Si D es una matriz diagonal y todos los términos de la diagonal son distintos, ¿qué matrices A son conmutables con D (esto es, $AD = DA$)?
*14. Probar que todas las matrices conmutables con la matriz D del Ejerc. 1 pueden expresarse en la forma $aI + bD$.
15. Si A es una matriz $n \times n$, probar que el conjunto $C(A)$ de todas las matrices $n \times n$ conmutables con A es cerrado para la adición y la multiplicación.
*16. Probar que cada matriz $n \times n$, A, satisface a una ecuación de la forma

$$A^m + c_{m-1}A^{m-1} + \ldots + c_1A + c_0I = 0, \qquad m \leq n^2.$$

*17. a) Sea $A = \| a_{ij} \|$ una matriz de números reales en la cual todos sus elementos están acotados, o sea, que para algún M fijo, cada $|a_{ij}| < M$. Probar que los elementos de A^k están acotados por $n^{k-1}M^k$.
b) Demostrar que la serie $I + A + A^2/2! + A^3/3! + \ldots$ es siempre convergente. (Esto puede utilizarse para definir la función exponencial e^A de la matriz A.)

En los ejerc. 18-22, seguimos la notación del anterior Apéndice.

18. a) Probar que D es lineal; b) Demostrar que $e^D = J$.
19. Probar que $D_xD_y = D_yD_x$.
*20. a) Simplificar $xD - Dx$, $x\Delta - \Delta x$, $x\Delta^2 - \Delta^2 x$.
b) Simplificar $x^iD^j - D^jx^i$, $x^i\Delta^j - \Delta^jx^i$.
*21. Definido el operador ∇^2 de Laplace por $\nabla^2 = D_x^2 + D_y^2$, hallar $x\nabla^2 - \nabla^2 x$, $y(\nabla^2)^2 - (\nabla^2)^2 y$, $\nabla^2(x^2+y^2) - (x^2+y^2)\nabla^2$.
*22. Desarrollar $\Delta^n = (J - I)^n$ por un «teorema del binomio».

3. Matrices rectangulares

Una matriz rectangular $m \times n$, A, sobre un campo F es una disposición rectangular $\| a_{ij} \|$ de números de F, que se distribuyen en m filas ($i = 1, \ldots, m$) y n columnas ($j = 1, \ldots, n$). La suma de dos matrices $m \times n$ y el producto de una matriz $m \times n$ por un escalar, se definen como en § 2; de aquí resulta que el conjunto de las matrices $m \times n$ constituye un espacio vectorial (cfr. Teor. 5).

Una matriz $m \times n$, $A = \| a_{ij} \|$, y una matriz $n \times r$, $B = \| b_{jk} \|$, con *el mismo* valor de n, pueden multiplicarse: el producto $AB = \| c_{ik} \|$ es una matriz $m \times r$ cuyos elementos están dados por la fórmula $c_{ik} = \sum_j a_{ij}b_{jk}$ (en esta suma, j varía desde 1 hasta n). Este producto de «filas por columnas» no podría definirse sin que cada fila de A

tuviese la misma longitud que cada columna de B; por esto hemos supuesto que el número de filas de B fuese igual al de columna de A. Así, si $m=1$, $n=2$, $r=3$,

$$(x_1, x_2)\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = (x_1a_{11}+x_2a_{21},\ x_1a_{12}+x_2a_{22},\ x_1a_{13}+x_2a_{23}).$$

Las leyes algebraicas explicadas antes para las matrices cuadradas, son también válidas para las matrices rectangulares, con tal que éstas tengan las dimensiones necesarias para que los productos sean posibles.

Por ejemplo: la matriz idéntica $m \times m$, I_m, y la matriz idéntica $n \times n$, I_n, satisfacen las igualdades

(24) $$I_mA = A = AI_n \qquad (A \text{ es } m \times n).$$

Se deduce que la multiplicación de matrices es bilineal, como en (22) y (23). La ley asociativa es

(25) $$A(BC)=(AB)C \quad (A \text{ es } m \times n,\ B \text{ es } n \times r,\ C \text{ es } r \times s).$$

Una vez más, el mejor método de prueba es acudir a la interpretación de las matrices rectangulares como transformaciones. La disposición rectangular $m \times n$, $\|a_{ij}\|$, da un sistema correspondiente de ecuaciones lineales homogéneas

(26) $$y_j = x_1a_{1j}+x_2a_{2j}+\dots+x_ma_{mj} \qquad (j=1, \dots, n),$$

que expresa las n variables y, en función de m variables x. Geométricamente, las x_i pueden ser coordenadas de un vector ξ relativas a cierta base de un espacio m-dimensional V, mientras las y_j son coordenadas de un vector η en *otro* espacio W de n dimensiones. Las ecuaciones (26) representan, pues, una transformación $\xi T=\eta$ del espacio m-dimensional V en el n-dimensional W, y ésta es lineal, pues $(c\xi_1+d\xi_2)T=c(\xi_1T)+d(\xi_2T)$. Recíprocamente, toda transformación lineal de V_m en W_n puede representarse, respecto a bases dadas de V_m y W_n, por ecuaciones lineales homogéneas (26), deducidas de una matriz $m \times n$, A.

La transformación T de V_m en W_n puede multiplicarse por una segunda transformación lineal, U, sólo si U, operando sobre el espacio imagen W_n de T, lo transforma en un tercer espacio S_r. El producto TU transforma entonces V_m en S_r y se representa por una

matriz $m \times r$, que es el producto AB de la matriz $m \times n$ representante de T por la matriz $n \times r$ representante de U (Teor. 4). Nuevamente resulta que el producto existe sólo cuando los dos factores tienen «común» la dimensión n. La ley asociativa para estas transformaciones implica la ley asociativa (25).

La *transpuesta* A' de una matriz $m \times n$, A, es una matriz $n \times m$, A', con elementos $a_{ij}' = a_{ji}$ ($i = 1, \ldots, n$; $j = 1, \ldots, m$). La fila i de la matriz transpuesta A' es la columna i de la matriz original A, y viceversa. Se puede también obtener A' por reflexión de A en su diagonal principal, o sea la de elementos a_{ii}. Para calcular la traspuesta C' de un producto $AB = C$, observemos que

$$c_{ik}' = c_{ki} = \sum_j a_{kj} b_{ji} = \sum_j b_{ji} a_{kj} = \sum_j b_{ij}' a_{jk}' ;$$

el resultado es el elemento (i, k) del producto $B'A'$ (obsérvese el cambio del orden). Esto prueba la primera de las leyes

(27) $$(AB)' = B'A', \quad (A + B)' = A' + B', \quad (cA)' = cA'.$$

La correspondencia $A \leftrightarrow A'$ conserva, pues, las sumas e invierte el orden de los productos, por lo cual es llamado, a veces, *antiautomorfismo*; puesto que $(A')' = A$, este antiautomorfismo se llama *involutivo*.

El uso sistemático de matrices rectangulares tiene ventajas formales diversas. Las coordenadas de un vector ξ, respecto a una base dada, en un espacio n-dimensional V, escritas en una fila, aparecen así:

(28) $$X = (x_1, \ldots, x_n), \quad X, \text{ matriz de una fila.}$$

Esto puede considerarse como un vector en $V_n(F)$, o como una matriz de una fila. Con tal matriz, pueden formarse matrices productos, como en el caso particular calculado en la fórmula inmediata anterior a (24). Las ecuaciones de una transformación lineal $y_j = \Sigma x_i a_{ij}$ tiene la forma de un producto matricial: la matriz de una fila Y se obtiene como producto de la matriz de una fila X por la matriz A. La transformación lineal con matriz A (respecto a una base dada) es, por lo tanto, representada en la forma reducida

(29) $$Y = XA \quad (X, Y, \text{ matrices de una fila}).$$

La traspuesta X' de una matriz de una fila, es una matriz de na columna. La ecuación (29) de una transformación lineal podía ambién escribirse en función de matrices de una columna, pues omando la traspuesta de cada miembro de (29) tenemos $Y'=A'X'$. El producto interno $x_1y_1+\ldots+x_ny_n$ de dos vectores (Cap. VII, § 7) s la matriz producto de una matriz de la fila X por la matriz de na columna Y', así que

$$(30)\qquad (X,\ Y)=XY' \qquad (X,\ Y,\ \text{matrices de una fila}).$$

La multiplicación de filas por columnas de dos matrices A y B s, por lo que precede, una multiplicación de la fila i de la matriz A or la columna k de la matriz B, así que la definición de producto uede formularse

$$(31)\qquad AB=\| c_{ik} \| \qquad \text{siendo } c_{ik}=A_iB^{(k)},$$

onde hemos empleado la notación

$$(32)\qquad A_i=\text{fila } i \text{ de } A, \qquad B^{(k)}=\text{columna } k \text{ de } B.$$

En toda la fila i $(c_{i1}, \ldots, c_{in})$ del producto AB, intervienen solanente la fila i de A y todas las columnas de B, o sea que es la mariz producto de A_i por toda la B. Similarmente, la columna k de AB se obtiene multiplicando la columna k de B por las filas de A.

En la notación de (32), estas reglas se formulan

$$(33)\qquad (AB)_i=A_iB, \qquad (AB)^{(k)}=AB^{(k)}.$$

Estas reglas pueden hacerse más sugestivas expresando a B como la fila de sus columnas, pues entonces

$$(34)\qquad A\cdot\| B^{(1)}\ B^{(2)}\ \ldots\ B^{(r)} \|=\| AB^{(1)}\ AB^{(2)}\ \ldots\ AB^{(r)} \|.$$

Estas columnas pueden también agruparse en conjuntos de columnas, formando submatrices. Así, una matriz 6×5, B, puede considerarse como una matriz 6×2, $D_1=\| B^{(1)}\ B^{(2)} \|$, adosada a una matriz 6×3, $D_2=\| B^{(3)}\ B^{(4)}\ B^{(5)} \|$, de modo que la matriz 6×5 será $B=\| D_1\ D_2 \|$. Por (34), la regla de multiplicación es ahora

$$(35)\qquad A\cdot\| D_1\ D_2 \|=\| AD_1\ AD_2 \|;\ D_1,\ D_2, \text{ bloques de } n \text{ filas}.$$

Estas dos fórmulas son casos particulares del método para multiplicar matrices que han sido subdivididas en «bloques» o submatrices. Es conveniente presentar otros ejemplos de este método:

$$\left\|\begin{array}{cccc|cccc} a_{11} & \dots & a_{1s} & & a_{1,s+1} & \dots & a_{1n} \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ a_{m1} & \dots & a_{ms} & & a_{m,s+1} & \dots & a_{mn} \end{array}\right\| \qquad \left\|\begin{array}{ccc} b_{11} & \dots & b_{1r} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ b_{s1} & \dots & b_{sr} \\ \hline b_{s+1,1} & \dots & b_{s+1,r} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ b_{n1} & \dots & b_{nr} \end{array}\right\| \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\} N_1 \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\} N_2 \end{array}$$

$$\underbrace{\qquad\qquad}_{M_1} \qquad \underbrace{\qquad\qquad}_{M_2}$$

Supongamos las n columnas de una matriz A descompuestas en dos grupos, el primero formado por las s columnas de una submatriz M_1 y el segundo por las restantes $n-s$ columnas, que formarán la submatriz M_2. Hagamos una subdivisión análoga con las filas de la matriz B tal, que B aparezca como una matriz $s \times r$, N_1, sobre una matriz $(n-s) \times r$, N_2. La fórmula del producto para $AB=C$ se subdivide en dos partes correspondientes

$$(36) \qquad c_{ik} = (a_{i1}b_{1k} + \dots + a_{is}b_{sk}) + (a_{i,s+1}b_{s+1,k} + \dots + a_{in}b_{nk}).$$

En el primer paréntesis aparecen sólo la fila i de la primera submatriz M_1 de A y la columna k de la submatriz superior N_1 de B. Por eso, este primer paréntesis es exactamente d_{ik}, el elemento de la fila i y columna k del producto M_1N_1. Análogamente, el segundo paréntesis de (36) es el término d^*_{ik} del producto M_2N_2. Por esto, $c_{ik}=d_{ik}+d^*_{ik}$, así que el producto total AB es la matriz suma $M_1N_1+M_2N_2$. Resulta, pues,

$$(37) \qquad \| M_1 \;\; M_2 \| \cdot \left\| \begin{array}{c} N_1 \\ N_2 \end{array} \right\| = M_1N_1 + M_2N_2.$$

Esta fórmula es una multiplicación de *filas por columnas* de *bloques* (submatrices), análoga en la multiplicación de filas por columnas de elementos. Un resultado semejante obtenemos para cualquier subdivisión de las columnas de A, con una correspondiente subdivisión de las filas de B. Cuando filas y columnas están simul-

táneamente subdivididas, la regla para la multiplicación es una combinación de (37) y (35),

$$(38)\quad \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \cdot \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} M_{11}N_{11}+M_{12}N_{21} & M_{11}N_{12}+M_{12}N_{22} \\ M_{21}N_{11}+M_{22}N_{21} & M_{21}N_{12}+M_{22}N_{22} \end{pmatrix}.$$

Esto supone que en la subdivisión hecha, el número de columnas de M_{11} es igual al número de filas de N_{11}. Esta regla (38) es exactamente la regla de multiplicación de matrices 2×2, establecida en § 2 (15), con la diferencia de que los elementos M_{ij} y N_{ij} son ahora submatrices o bloques y no escalares. Concluimos, pues, que la multiplicación de matrices divididas convenientemente en bloques se efectúa lo mismo que la multiplicación ordinaria de matrices.

EJERCICIOS

1. Sean

$$A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix},\quad B=\begin{pmatrix} i & 0 & -i \\ 0 & 1 & 1+i \end{pmatrix},\quad X=(1, -1),\quad Y=(i, 0).$$

a) Hallar XA, XB, YA, YB;
b) Hallar $3A-4B$, $A+(1+i)B$, $(X-(1+i)Y)(iA+5B)$;
c) Hallar BA', AB', XAB', $BA'Y$.

2. Una transformación T hace corresponder al $(1, 1)$ el $(0, 1, 2)$ y al $(-1, 1)$ el $(2, 1, 0)$. ¿Qué matriz representa a T?
3. Demostrar que $(A+B)'=A'+B'$, $(cA)'=cA'$.
4. Hallar AB, BA, AC y BC, siendo

$$A=\begin{pmatrix} 2 & 3 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},\quad B=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix},\quad C=\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

5. Sea I^* la matriz $(r+n)\times n$ que se forma colocando una matriz identidad $n\times n$, sobre una matriz $r\times n$ de ceros. ¿Cuál es el efecto de multiplicar cualquier matriz $n\times(r+n)$ por I^*?
6. Demostrar la regla (38) para la multiplicación por submatrices.

4. Inversas

Las transformaciones lineales pueden dividirse en dos categorías, según que sean biunívocas o no lo sean. Por el Teorema 1 del Capítulo VI, sabemos que la transformación $\eta=\xi T$ de un espacio

vectorial V es una correspondencia biunívoca $\xi \to \eta$ de V consigo mismo si, y sólo si, existe una transformación inversa T^{-1} definida formalmente.

DEFINICIÓN. *Una transformación lineal* T *de un espacio vectorial* V *se llamará «no-singular» o «regular», si existe una transformación* T^{-1} *de* V *tal, que* $TT^{-1}=T^{-1}T=I$, *siendo* I *la transformación idéntica. En el caso contrario, la transformación* T *se llamará singular.*

Una transformación no singular T es una correspondencia biunívoca de V con V, que conserva las operaciones algebraicas de suma y producto escalar; es, pues, un isomorfismo del espacio vectorial V consigo mismo. Por esto, a una transformación lineal no singular de V se la puede llamar un «automorfismo» de V.

El método más directo para establecer las propiedades principales de las transformaciones lineales, singulares y no singulares, es aplicar la teoría de la independencia lineal, desarrollada en el Capítulo VII, utilizando una base fija $\epsilon_1, \ldots, \epsilon_n$ para el espacio vectorial V sobre el cual opera la transformación dada T. Respecto a esta base, T se representa como en (29) por las ecuaciones lineales $Y=XA$, con una matriz cuadrada $A=\| a_{ij} \|$.

Designemos por $\alpha_1=\epsilon_1 T$, $\alpha_2=\epsilon_2 T$, ..., $\alpha_n=\epsilon_n T$ los transformados por T de los vectores base. De las ecuaciones (9) de la transformación, se deduce inmediatamente (*) que las coordenadas del k-ésimo vector transformado $\alpha_k=\epsilon_k T$ son los elementos de la fila k, A_k, de la matriz A.

Por «resultante» de T entenderemos el conjunto de todos los vectores ξT que se obtienen aplicando T. Puesto que

$$(39) \qquad \xi T=(x_1\epsilon_1+\ldots+x_n\epsilon_n)T=x_1\alpha_1+\ldots+x_n\alpha_n,$$

el resultante de T está constituído por todas las combinaciones lineales de los vectores α_k, o sea, que es el subespacio engendrado por los α_k. Evidentemente habrá que distinguir dos casos posibles, según que los vectores α_k (las filas de A) sean o no linealmente independientes.

(*) Inversamente, como los vectores base $\epsilon_1, \ldots, \epsilon_n$ tienen por coordenadas las correspondientes filas $I_1, \ldots, I_n$ de la matriz idéntica I, la ecuación $Y=XA$ y la regla (33) dan para las coordenadas de las α_k la expresión $I_k A=A_k$.

Caso I. Si los α_k son linealmente independientes, constituyen una *base* de V, por el Teor. 8, Cap. VII. En este caso, cada uno de los vectores η de V puede escribirse de un modo único como una combinación lineal de los α_k,

$$\eta = x_1\alpha_1 + \ldots + x_n\alpha_n = x_1(\epsilon_1 T) + \ldots + x_n(\epsilon_n T),$$

y por lo tanto, η es el transformado (39) de un véctor, y sólo de uno, a saber, de $\xi = x_1\epsilon_1 + \ldots + x_n\epsilon_n$. T es, por consiguiente, una correspondencia biunívoca, y la transformación inversa $\eta \rightarrow \xi$ está dada por la fórmula

$$(40) \qquad (x_1\alpha_1 + \ldots + x_n\alpha_n)T^{-1} = x_1\epsilon_1 + \ldots + x_n\epsilon_n.$$

Dicho de otro modo, las coordenadas $x_1, \ldots, x_n$ del vector η respecto a la base $\alpha_1, \ldots, \alpha_n$, son también las coordenadas del vector transformado $\eta T^{-1} = \xi$ respecto a la base original. Por otra parte, respecto a *cualquier* base $\beta_1, \ldots, \beta_n$, las sumas y los productos por escalares se calculan efectuando las operaciones correspondientes con las coordenadas, según las fórmulas $\Sigma x_i\beta_i + \Sigma y_i\beta_i = \Sigma(x_i + y_i)\beta_i$ y $c(\Sigma x_i\beta_i) = \Sigma(cx_i)\beta_i$. En consecuencia, la transformación $\eta T^{-1} = \xi$ de (40) transforma sumas en sumas y productos por escalares en productos por escalares, y será, por consiguiente, una transformación lineal. En resumen, si las filas de A son linealmente independientes, T tiene una *inversa lineal* T^{-1} y es regular.

Caso II. Si las α_k son linealmente dependientes, la cosa es muy distinta. En este caso, por definición, habrá escalares c_k no todos cero, tales que

$$c_1\alpha_1 + \ldots + c_n\alpha_n = (c_1\epsilon_1 + \ldots + c_n\epsilon_n)T = O,$$

donde $\gamma = c_1\epsilon_1 + \ldots + c_n\epsilon_n \neq O$. Entonces, para cualquier transformación U, lineal o no, $\gamma(TU) = (\gamma T)U = (O)U = (OT)U = O(TU)$; en consecuencia, una de las relaciones $\gamma(TU) = \gamma$, $O(TU) = O$ debe ser falsa, lo que nos dice que TU no puede ser la transformación idéntica I. Por consiguiente, U no será jamás una inversa por la derecha de T.

Además, en el Caso II, el resultante de T está engendrado por los n vectores dependientes $\alpha_1, \ldots, \alpha_n$; luego, por el Teorema 7 del Capítulo VII, tiene una dimensión menor que n. Por consiguiente, del resultante de T debe excluirse algún elemento β de V. Para

este β y cualquier U, luego [illegible] $\beta U T$ es del resultado de T, y no puede ser igual a β. Concluimos que UT no puede ser la identidad I, es decir, U no es inverso de T por la izquierda. En resumen, si las filas de A son linealmente dependientes, T no tiene ni inverso a la derecha ni inverso a la izquierda, sea éste lineal o no, y por lo tanto, T es singular.

Así, el Caso I es el de transformación regular, y el Caso II, el de transformación singular, independientemente de la elección de base en V. A causa del isomorfismo multiplicativo entre transformaciones lineales y matrices, vemos en el Caso I que A tiene una matriz inversa A^{-1} tal, que

$$(41) \qquad A^{-1}A = I = AA^{-1} \qquad (A,\ A^{-1},\ I,\ \text{matrices } n \times n),$$

mientras que en el Caso II, A no tiene inversa por la derecha ni por la izquierda. Esto nos lleva a llamar «no singular» o «regular» a una matriz si, y sólo si, tiene una inversa A^{-1} que satisface a las condiciones (41). Si A tiene una inversa, también la tiene su transpuesta, como se ve transponiendo los dos miembros de (41), con lo que obtenemos $A'(A^{-1})' = I = (A^{-1})'A'$, y por tanto,

$$(42) \qquad (A^{-1})' = (A')^{-1}.$$

Además, A' tendrá un inverso siempre que A lo tenga, así que lo demostrado sobre la independencia de las filas de A se aplica igualmente a la independencia de columnas. Los dos casos pueden reunirse como sigue:

TEOREMA 7. *Si una transformación lineal* T *no es singular, tiene una transformación lineal inversa, y las filas (y columnas) de cualquier matriz* A *que represente a* T *son linealmente independientes. Si* T *es singular, no tiene inversa ni por la izquierda ni por la derecha, y las filas, y también las columnas, de cualquier matriz* A *representante de* T *son linealmente dependientes.*

COROLARIO 1. *Una transformación lineal* T *sobre un espacio con base* $\epsilon_1, \ldots, \epsilon_n$ *es no singular si, y sólo si, los transformados* ϵ_1T, …, ϵ_nT *de los vectores base son linealmente independientes.*

COROLARIO 2. *Si el producto* TU *de dos transformaciones lineales es la identidad, entonces* T *y* U *son ambas biunívocas,* T [illegible] U [illegible] T^{-1} *y* U [illegible] T [illegible]

Demostración. $TU = I$ significa que U es una inversa a la derecha de T; por esto, y por el teorema, T no puede ser singular. Entonces sabemos que su inversa (a la derecha) $U = T^{-1}$ está determinada unívocamente.

COROLARIO 3. *Una matriz cuadrada A es no singular si sus filas son linealmente independientes y sólo en este caso.*

Demostración. Cada matriz A determina una transformación lineal correspondiente $Y = XA$, que es no singular si la matriz A es regular (y viceversa). El resultado se deduce entonces por las condiciones del teorema. Una aplicación análoga del Corolario 2 demostrará:

COROLARIO 4. *Toda inversa de una matriz por la izquierda, es también inversa por la derecha y recíprocamente.*

Si dos matrices A y B tienen inversas, también lo tiene su producto:

$$(AB)^{-1} = B^{-1}A^{-1} \qquad \text{(¡nótese el orden!)} \tag{43}$$

Una transformación regular T de V_n es un automorfismo, luego transformará líneas en líneas, planos en planos, etc. Dado un subespacio S de V_n, el conjunto de todos los transformados ξT de los vectores ξ de S es también un subespacio S', pues cada combinación lineal $c(\xi T) + d(\eta T) = (c\xi + d\eta)T$ pertenece al conjunto S'. Este espacio S' puede ser llamado el transformado $S' = ST$ de S. Su dimensión es siempre igual a la de S.

TEOREMA 8. *Una transformación lineal no singular transforma cada subespacio en otro subespacio de la misma dimensión.*

Demostración. Si S tiene dimensión d, contiene d vectores independientes $\alpha_1, \ldots, \alpha_d$, que forman parte de una base de V_n. Sus imágenes $\alpha_1 T, \ldots, \alpha_d T$ están en S' y son independientes, por el Corolario 1 del Teorema 7; luego el espacio transformado tiene dimensión $d' \geq d$. El mismo razonamiento para T^{-1} prueba que $d' \leq d$; luego $d = d'$, como queríamos demostrar.

El cálculo de la matriz inversa de otra regular dada puede hacerse resolviendo en cada caso un sistema conveniente de ecuacio-

nes lineales. Escribamos las coordenadas de los vectores de la base como sigue :

(44) $I_1=(1, 0, 0, \dots, 0), I_2=(0, 1, 0, \dots, 0), \dots, I_n=(0, 0, \dots, 0, 1)$.

Entonces, en una matriz dada $A=\| a_{ij} \|$, cada fila A_i resulta una combinación lineal $A_i=\sum_j a_{ij}I_j$ de los vectores base. Estas ecuaciones pueden resolverse para las «incógnitas» I_j en función de las A_i ; el resultado dará una expresión lineal para cada I_j,

(45) $$I_j=c_{j1}A_1+\dots+c_{jn}A_n=\sum_k c_{jk}A_k.$$

Esta fórmula nos dice que la fila I_j de lugar j en la matriz identidad, es la fila j de una matriz $C=\| c_{ij} \|$ multiplicada por una matriz A, que ha sido descompuesta en sus filas A_k. En otras palabras, $I=CA$, o sea que C es la matriz inversa pedida : $C=A^{-1}$. En Capítulo X, § 3, daremos otra construcción para A^{-1}.

EJEMPLO. Para hallar la inversa de la matriz que aquí damos, se escriben sus filas como $A_1=I_1+2I_2-2I_3$, $A_2=-I_1+3I_2$, $A_3=-2I_2+I_3$.

$$\begin{pmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

Estas tres ecuaciones simultáneas tienen una solución $I_1=3A_1+2A_2+6A_3$, $I_2=A_1+A_2+2A_3$, $I_3=2A_1+2A_2+5A_3$.

Los coeficientes c_{jk} de estas combinaciones lineales dan la matriz inversa ; efectivamente, se tiene

$$\begin{pmatrix} 3 & 2 & 6 \\ 1 & 1 & 2 \\ 2 & 2 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & -2 \\ -1 & 3 & 0 \\ 0 & -2 & 1 \end{pmatrix}=\begin{pmatrix} 3-2 & 6+6-12 & -6+6 \\ 1-1 & 2+3-4 & -2+2 \\ 2-2 & 4+6-10 & -4+5 \end{pmatrix}=$$
$$=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

EJERCICIOS

1. Hallar las inversas de las matrices A, B, C, D, del Ejerc. 1, § 2.
2. a) Probar que $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es no singular si, y sólo si, $ad-bc \neq 0$.

   b) Hallar una fórmula para las inversas de una matriz 2×2 no singular.
3. Hallar las inversas de las transformaciones lineales R_θ, U_b, S_k, H_a de § 1.
4. Hallar las inversas (si existen) de las transformaciones lineales del Ejercicio 4, § 1.

5. a) Si $\theta=45°$, calcular la matriz de la transformación $R_\theta^{-1}U_b R_\theta$ (ver § 1).
   b) Describir geométricamente el efecto de esta transformación.
   c) Hacer lo mismo para $R_\theta^{-1}H_a R_\theta$ (con $\theta=45°$).
6. Si A satisface a $A^2-A+I=0$, probar que A^{-1} existe y es igual a $I-A$.
7. Hallar las inversas de las matrices E_1, E_2 y E_3 del Ejerc. 9, § 2.
8. Hallar las inversas de las matrices A y B del Ejerc. 4, § 3. (*Sugerencia:* Descomponer en bloques.)
9. Se llama matriz *triangular* aquella en la que todos los elementos debajo de la diagonal principal son ceros.
   a) Obtener una fórmula para la inversa de una matriz 2×2 triangular.
   b) Lo mismo para una matriz 3×3 triangular. (*Sugerencia:* Ensayar una inversa triangular.)
   c) Probar que toda matriz triangular con ningún término nulo en la diagonal principal tiene una inversa triangular.
10. Dados A, B, A^{-1}, B^{-1} y C, hallar las inversas multiplicativas de

    a) $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, b) $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$. c) $\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$.
11. Probar que las matrices no singulares forman grupo respecto a la multiplicación de matrices.
12. Probar que una transformación lineal T no es singular si, y sólo si, $\xi T=O$ implica $\xi=O$.
13. Si un producto AB no es singular, probar que ambos factores A y B son regulares.
*14. Probar directamente, a partir de las definiciones, que la inversa de una transformación lineal es necesariamente lineal.
*15. Probar, sin apelar a las transformaciones lineales, que una matriz A tiene una inversa a la izquierda si, y sólo si, sus filas son linealmente independientes. [*Sugerencia:* Deducirlo de (45).]
*16. Probar el corolario 4 sin apelar a las transformaciones lineales. (*Sugerencia:* Demostrar que la inversa por la izquierda dada, tiene sus filas linealmente independientes.)

5. Cuaternios

Las leyes algebraicas válidas para matrices cuadradas se aplican a otros sistemas algebraicos, tales como los cuaternios de Hamilton. Éstos son cantidades que resultan de adjuntar a los ordinarios números complejos con $\pm i$, otra cantidad, j, que representa otra raíz cuadrada de -1. No es posible construir un campo que contenga a ambos elementos i y j, pues en un campo una ecuación cuadrática, $x^2=-1$, puede tener a lo más las dos raíces $\pm i$. Por esto, no impondremos que la multiplicación de j por un número complejo sea conmutativa, sino que seguirá estas reglas:

$$(46) \qquad j^2=-1, \qquad (a+bi)j=j(a-bi).$$

En función de los números complejos $u=a+bi$ y de sus conjugados $u^*=a-bi$, la segunda regla se formula: $uj=ju^*$. En particular, $ij=-ji$. Si llamamos $k=ij$, la regla (46) da una tabla de multiplicación para 1, i, j y k:

$$(47) \qquad i^2=j^2=k^2=-1, \quad ij=-ji=k, \quad jk=-kj=i, \quad ki=-ik=j.$$

Llamaremos *cuaternios* a las cantidades x, y, de la forma

$$(48) \qquad x=x_0+x_1i+x_2j+x_3k, \qquad y=y_0+y_1i+y_2j+y_3k,$$

donde los coeficientes x_i, y_i, son números reales; x e y pueden escribirse en forma más breve como

$$(49) \qquad x=u_1+u_2j, \qquad y=v_1+v_2j,$$

donde $u_1=x_0+x_1i$, $u_2=x_2+x_3i$, v_1 y v_2 son números complejos. Dos cuaternios son iguales si, y sólo si, son formalmente idénticos; así, en (48), $x=y$ significa que $x_i=y_i$ para $i=0, 1, 2, 3$.

Las operaciones algebraicas sobre los cuaternios x e y son tres:

Multiplicación escalar: Para multiplicar x por el número real c se multiplica cada componente x_i por c, o sea,

$$(50) \qquad cx=cx_0+cx_1i+cx_2j+cx_3k.$$

Adición: Se obtiene la suma de dos cuaternios por adición de componentes correspondientes, es decir,

$$(51) \qquad x+y=(x_0+y_0)+(x_1+y_1)i+(x_2+y_2)j+(x_3+y_3)k.$$

Multiplicación: El producto de x por y se obtiene multiplicando las dos expresiones (48) usando la ley distributiva y la tabla de multiplicación (47) para los generadores i, j, k. Esta definición del producto puede expresarse mediante los coeficientes complejos de (49) así:

$$(52) \qquad (u_1+u_2j)\,(v_1+v_2j)=(u_1v_1-u_2v_2^*)+(u_1v_2+u_2v_1^*)j.$$

El resultado de efectuar estas tres operaciones entre x e y es siempre otro cuaternio.

La multiplicación escalar y la adición están establecidas como si x e y fuesen vectores con las componentes respectivas x_i e y_i.

Por esto, los cuaternios x forman un espacio lineal vectorial Q sobre el campo de los números reales, con una base $1, i, j, k$ de cuatro elementos linealmente independientes. Por todas estas definiciones, el producto es distributivo para la suma, mientras que, por la definición (52), se puede comprobar que el producto es también bilineal y asociativo. Existe un elemento unidad $1=1+0i+0j+0k$. Existe también un inverso a la derecha para todo $x\neq 0$. Para hallarlo se acude al llamado «conjugado» $\bar{x}=x_0-x_1i-x_2j-x_3k=$ $=u_1^*-u_2j$, y a la «norma» $N(x)=x\bar{x}$. Por (52),

$$x\bar{x}=(u_1+u_2j)\,(u_1^*-u_2j)=u_1u_1^*+u_2u_2^*$$

Pero el producto uu^* de un complejo por su conjugado es real y no negativo. Así, $u_1u_1^*+u_2u_2^*$ es un escalar $c\neq 0$, y $(1/c)\bar{x}$ es el inverso a la derecha de x.

Estas leyes algebraicas para cuaternios se comprenden en el

TEOREMA 9. *Los cuaternios satisfacen todos los postulados para un campo, excepto la ley conmutativa de la multiplicación.*

Todos los cuaternios x satisfacen a una ecuación cuadrática $f(t)=0$, con raíces x y $\bar{x}$ y con coeficientes reales, pues

$$(t-x)\,(t-\bar{x})=t^2-(x+\bar{x})t+x\bar{x}=t^2-2x_0t+(u_1u_1^*+u_2u_2^*).$$

Un cuaternio «puro» es, por definición, el que tiene el primer coeficiente cero. Según (48), el producto de dos cuaternios puros x e y, es

$$(x_1i+x_2j+x_3k)\,(y_1i+y_2j+y_3k)=-x_1y_1-x_2y_2-x_3y_3+$$
$$+(x_2y_3-x_3y_2)i+(x_3y_1-x_1y_3)j+(x_1y_2-x_2y_1)k.$$

La parte «pura» de este resultado es llamada en física el «producto vectorial» (o producto externo) :

$$(x_1,\ x_2,\ x_3)\times(y_1,\ y_2,\ y_3)=(x_2y_3-x_3y_2,\ x_3y_1-x_1y_3,\ x_1y_2-x_2y_1).$$

En la media centuria de 1850-1900, gran parte del actual análisis vectorial tridimensional solía expresarse en el lenguaje de los cuaternios.

EJERCICIOS

1. Resolver $xc=d$ para: a) $c=i$, $d=1+j$; b) $c=2+j$, $d=3+k$.
2. a) Demostrar que $x^2=-1$ tiene como soluciones una infinidad de cuaternios x.
   b) Mostrar que esto no es contradictorio con el Teor. 3 del Cap. IV. sobre el número de raíces de un polinomio.
3. Sean $a=1+i+j$, $b=1+j+k$.
   a) Hallar $a+b$, ab, $a-b$, $ia-2b$, $\bar{a}$, $a\bar{a}$.
   b) Resolver $ax=b$, $xa=b$, $x^2=b$, $bx+(2j+k)=a$.
4. Deducir de (47) la tabla de multiplicar (52).
5. a) Demostrar que la norma $N(x)=x\bar{x}$ de x es $x_0^2+x_1^2+x_2^2+x_3^2$.
   b) Demostrar que $\overline{xy}=\bar{y}\cdot\bar{x}$.
   c) Demostrar que $N(x)N(y)=N(xy)$.
6. Demostrar que el inverso por la derecha de un cuaternio es siempre un inverso por la izquierda.
7. Probar que la solución de una ecuación de cuaternios $xa=b$ está unívocamente determinada si $a\neq 0$.
8. Si un cuaternio x satisface a una ecuación cuadrática $x^2+a_0x+b_0=0$, con coeficientes reales a_0 y b_0, probar que todo cuaternio $q^{-1}xq$ satisface a la misma ecuación cuadrática (si $q\neq 0$).
9. En el álgebra de cuaternios, probar que los elementos ± 1, $\pm i$, $\pm j$, $\pm k$, forman un grupo multiplicativo. (Este grupo, que podía ser definido directamente, es el llamado *grupo cuaternio*.)
10. Hallar en el grupo cuaternio los subgrupos normales de índice 2.
11. Probar que la multiplicación de cuaternios es asociativa.
12. a) Los cuaternios $x_0+x_1i+x_2j+x_3k$ con coeficientes racionales x_i, ¿forman un álgebra de división (ver § 6) sobre el campo de los números racionales? Probar la proposición.
   *b) Considerar la misma cuestión para los números complejos x_i como coeficientes. (*Nota:* En este problema, el número complejo $\sqrt{-1}$ como coeficiente no es el mismo que la i del cuaternio.)
13. Demostrar que el «producto vectorial» de dos vectores no es asociativo.
14. Si los enteros a y b son ambos suma de cuatro cuadrados enteros, demostrar que el producto ab es también una suma de cuatro cuadrados. (*Sugerencia:* Atender al Ejerc. 5.)
15. Deducir todas las reglas (47) a partir de $i^2=j^2=k^2=ijk=-1$.

6. Álgebras lineales

Se llama sistema de números hipercomplejos (o álgebra lineal) a un sistema algebraico que satisface las leyes formales de la adición, multiplicación escalar y multiplicación, válidas para matrices cuadradas. Los escalares que se usan pueden ser números reales, como en el caso de los cuaternios, o elementos de un campo F, como para la llamada «álgebra de las matrices» $M_n(F)$, constituída por todas las matrices $n\times n$ con elementos de F.

DEFINICIÓN. *Un álgebra lineal sobre un campo F es un conjunto $\mathfrak{A}$, el cual es un espacio vectorial sobre F de número finito de dimensiones, que admite una multiplicación asociativa y bilineal,*

(53) $$\alpha(\beta\gamma)=(\alpha\beta)\gamma, \qquad \textit{(asociativa)}$$

(54) $$\alpha(c\beta+d\gamma)=c(\alpha\beta)+d(\alpha\gamma), \quad (c\alpha+d\beta)\gamma=c(\alpha\gamma)+d(\beta\gamma), \qquad \textit{(bilineal)}$$

entendiendo que estas leyes se verifican para todas las escalares c *y* d *de* F *y todos los* α, β, γ *de* $\mathfrak{A}$. *El orden de* $\mathfrak{A}$ *es su dimensión considerado como espacio vectorial. Se dice que* $\mathfrak{A}$ *tiene un elemento unidad* ϵ *si éste es tal que* $\epsilon\alpha=\alpha=\alpha\epsilon$, *para todo* α *en* $\mathfrak{A}$. *El álgebra se llama un álgebra de división si, además de tener un elemento unidad* ϵ, *contiene para cada* $\alpha\neq 0$ *un* α^{-1} *para el cual* $\alpha^{-1}\alpha=\epsilon$.

Un célebre teorema de Frobenius (1878) establece que los cuaternios constituyen la única álgebra de división no conmutativa sobre el campo de los números reales.

EJEMPLO 1. Construyamos sobre los números reales un álgebra de «números duales» que tiene dos elementos básicos δ y ϵ, los cuales se multiplican según las reglas $\delta\epsilon=\epsilon\delta=\delta$, $\delta^2=0$, $\epsilon^2=\epsilon$. Con estas reglas podríamos hallar el producto de dos elementos cualesquiera de A, pues

$$(a\delta+b\epsilon)(c\delta+d\epsilon)=ac\delta^2+ad\delta\epsilon+bc\epsilon\delta+bd\epsilon^2=(ad+bc)\delta+bd\epsilon.$$

Los postulados requeridos, tales como la ley asociativa de la multiplicación, se verifican, según puede comprobarse. Este ejemplo, como el de los cuaternios, muestra que un álgebra queda definida dando una adecuada tabla de multiplicación para los elementos básicos.

EJEMPLO 2. El álgebra $M_n(F)$ compuesta por todas las matrices $n\times n$ sobre F, tiene como elementos base las matrices E_{ij}, que tienen un 1 en el elemento (i, j) y ceros todos los demás. La tabla de multiplicación para estas matrices es:

$$E_{ij}E_{jk}=E_{ik}, \quad E_{ij}E_{kl}=0 \quad (j\neq k).$$

EJEMPLO 3. Sea G un grupo finito, con elementos $a_1, \ldots, a_n$ y multiplicación $a_ia_j=a_k$. Si F es un campo, existe un álgebra

lineal $\mathfrak{A}$ sobre F, que tiene los elementos de G como base, y en la cual la multiplicación está determinada por la bilinealidad del grupo G,

$$(55)\qquad (x_1a_1+\ldots+x_na_n)(y_1a_1+\ldots+y_na_n)=\sum_{i,j}(x_iy_j)(a_ia_j).$$

Este álgebra se llama *álgebra del grupo* G sobre F.

EJEMPLO 4. El conjunto de todas las matrices $n\times n$ sobre F es un espacio lineal de dimensión n^2. Por esto, las sucesivas potencias I, A, A^2, A^3, ... de una matriz fija A, no pueden ser todas linealmente independientes. Si m es el más pequeño entero para el *cual las potencias* $I, A, \ldots, A^m$ *son dependientes, existe una ecuación*

$$(56)\qquad f(A)=A^m+c_{m-1}A^{m-1}+\ldots+c_1A+c_0I=0 \qquad (c_i \text{ en } F),$$

en la que no es cero el coeficiente c_m de A^m y por lo tanto puede tomarse igual a 1. Esta ecuación es la llamada ecuación *mínima* para A. Consideremos el conjunto $\mathfrak{S}$ de todas las matrices expresables *por polinomios en* A *con coeficientes* escalares. Puesto que una suma o producto de polinomios es un polinomio, el conjunto $\mathfrak{S}$ es la subálgebra de $M_n(F)$ engendrada por A. Puesto que la ecuación *mínima puede usarse para* expresar A^m *mediante potencias* de A con exponente menor, todos los polinomios en A pueden reducirse a polinomios de grado $m-1$ a lo más. Por lo tanto, $\mathfrak{S}$ es también un álgebra lineal de orden (dimensión) m sobre F.

EJEMPLO 5. El conjunto de todas las matrices $2n\times 2n$, con dos bloques $n\times n$ de ceros, situados encima a la derecha y debajo a la izquierda, *forman un álgebra*, que *es una* subálgebra de $M_n(F)$. Para comprobar esto necesitamos ver solamente que el producto de dos matrices con la estructura antedicha es otra matriz con la misma estructura. Esto *se comprueba por* las reglas del § 3 para multiplicación por bloques, pues

$$(57)\qquad \begin{pmatrix} A_1 & 0\\ 0 & A_2\end{pmatrix}\begin{pmatrix} B_1 & 0\\ 0 & B_2\end{pmatrix}=\begin{pmatrix} A_1B_1+0\cdot 0 & A_1\cdot 0+0\cdot B_2\\ 0\cdot B_1+A_2\cdot 0 & 0\cdot 0+A_2B_2\end{pmatrix}=$$
$$=\begin{pmatrix} A_1B_1 & 0\\ 0 & A_2B_2\end{pmatrix}.$$

Semejante resultado se obtendría con mayor número de bloques de diagonales, aunque no fuesen de igual dimensión.

DEFINICIÓN. *Dos álgebras* $\mathfrak{A}$ *y* $\mathfrak{A}'$ *sobre un mismo campo* F *son* isomorfas *(o equivalentes sobre* F*) si hay una correspondencia biunívoca* $\alpha \leftrightarrow \alpha'$ *entre* $\mathfrak{A}$ *y* $\mathfrak{A}'$ *que conserva las tres operaciones:*

$$(\alpha+\beta)'=\alpha'+\beta', \quad (c\alpha)'=c\alpha', \quad (\alpha\beta)'=\alpha'\beta'. \tag{58}$$

Nuestros postulados para el álgebra de matrices son completos, en el *siguiente sentido* (análogo al del Teorema de Cayley, Capítulo VI, § 5).

TEOREMA 10. *Toda álgebra lineal de orden* n *con elemento unidad, es isomorfa al álgebra de las matrices.*

Demostración. El álgebra $\mathfrak{A}$ es un espacio lineal de elementos ξ. Asociemos a cada elemento α de $\mathfrak{A}$ la transformación T que se obtiene multiplicando por él a la derecha, de modo que $\xi T=\xi\alpha$ para cualquier ξ de $\mathfrak{A}$. Puesto que la multiplicación es bilineal como en (54), T es una transformación lineal. Por otra parte, como existe la unidad ϵ, $\epsilon\alpha=\epsilon\beta$ implica $\alpha=\beta$, así que dos elementos distintos α y β inducen transformaciones distintas T y U. Además, los postulados del álgebra dan

$$\xi(\alpha+\beta)=\xi\alpha+\xi\beta, \quad \xi(c\alpha)=c(\xi\alpha), \quad \xi(\alpha\beta)=(\xi\alpha)\beta,$$

de modo que las transformaciones correspondientes son $\alpha+\beta \to T+U$, $c\alpha \to cT$, $\alpha\beta \to TU$. Esto significa que la *correspondencia* $\alpha \to T$ es un isomorfismo entre el álgebra dada y un álgebra de transformaciones lineales sobre $\mathfrak{A}$. Pero las transformaciones lineales están representadas isomórficamente por matrices, luego queda demostrado el teorema.

Para hacer explícitas tales matrices, elijamos una base $\epsilon_1, \ldots, \epsilon_n$ de $\mathfrak{A}$. La transformación T hará corresponder a cada ϵ_i algún elemento determinado $\epsilon_i\alpha=\sum_j c_{ij}\epsilon_j$. Luego

$$\left(\sum_i x_i\epsilon_i\right) T=\left(\sum_i x_i\epsilon_i\right)\alpha=\sum_{i,j} x_i c_{ij}\epsilon_j=\sum_j\left(\sum_i x_i c_{ij}\right)\epsilon_j.$$

Respecto a estas coordenadas, la transformación lineal T tiene, pues, las ecuaciones $y_j=\sum_i x_i c_{ij}$, con una matriz $C=\|c_{ij}\|$. El isomorfismo $\alpha \leftrightarrow C$ de $\mathfrak{A}$ con el álgebra de matrices C se llama *segunda representación regular de* $\mathfrak{A}$ (la «primera» resulta por premultiplicación con α y es un anti-isomorfismo).

Para dar un ejemplo, consideremos el campo de los números complejos como un álgebra (conmutativa) de orden 2 sobre los números reales, con base 1, i. Para $\alpha=i$, $1\cdot i=0+1\cdot i$, $i\cdot i=-1\cdot 1+0\cdot i$, así que la representación por matrices es

$$1\to\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\qquad i\to\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\qquad a+bi\to\begin{pmatrix}a & b\\ -b & a\end{pmatrix}.$$

La matriz correspondiente a i satisface, en efecto, a la ecuación $X^2=-I$. Estas matrices se emplean algunas veces para definir los números complejos.

EJERCICIOS

1. Demostrar que el elemento cero, 0, de un álgebra lineal satisface a $\xi\cdot 0=0=0\cdot\xi$ para todo ξ.
2. ¿Es el álgebra de los números duales un álgebra de división?

* 3. ¿Cuáles de los sistemas siguientes son álgebras lineales?
   a) Un espacio vectorial V_n con $\alpha\ \beta=0$ para todo α y β.
   b) Un espacio vectorial V_n con $\alpha\cdot\beta=\alpha$ para todo α y β.
   c) Todas las matrices $m\times n$ sobre un campo F.
   d) Todas las matrices $n\times n$ que tienen de ceros la diagonal principal (y los otros elementos son cualesquiera).
   e) Todas las matrices $n\times n$ triangulares (esto es, cuyos elementos debajo de la diagonal principal son todos nulos).
4. Hallar una tabla de multiplicación para el álgebra de cuaternios, si los elementos base son $\alpha=1-i$, $\beta=1+i$, $\gamma=j+k$, $\delta=j-k$.
5. Si P es una matriz no singular, probar que $A\to P^{-1}AP$ es un automorfismo del álgebra matricial completa.
6. a) Determinar las matrices correspondientes a 1, i, j y k, en la segunda representación regular del álgebra de cuaternios.
   b) Comprobar que estas matrices satisfacen la tabla de multiplicación dada para i, j, k en (47).
7. Mostrar las matrices representantes de los números duales δ y ϵ en la segunda representación regular.
8. Demostrar que las transformaciones lineales que representan números complejos son los productos de transformaciones de semejanza por rotaciones puras (§ 1).
9. a) Sea $\mathfrak{A}$ un álgebra lineal con base $\alpha_1, \ldots, \alpha_n$ sobre F. Probar que el producto de dos elementos cualesquiera de $\mathfrak{A}$ está unívocamente determinado cuando son conocidos (como en el caso de los cuaternios) los n^2 productos $\alpha_i\alpha_j=\sum_k c_{ij}^{(k)}\alpha_k$. Los escalares $c_{ij}^{(k)}$ son llamados «constantes de la multiplicación» y también «constantes de estructura».

b) En cualquier espacio vectorial, un producto está definido si se da el producto de dos elementos cualesquiera de las bases, α_i y α_j. Probar que este producto será asociativo si, y sólo si, $(\alpha_i\alpha_j)\alpha_k=\alpha_i(\alpha_j\alpha_k)$, para todo i, j, k.

c) Expresar esta condición asociativa en función de las «constantes de multiplicación» $c_{ij}^{(k)}$ de a).

d) Hallar una condición similar sobre las constantes $c_{ij}^{(k)}$ para un álgebra conmutativa.

10. Hallar las constantes de estructura para el álgebra de números complejos sobre los números reales, cuando las bases son $\alpha_1=1+i$, $\alpha_2=1-i$.

11. Probar que si A es una matriz $n\times n$ conmutativa con todas las matrices $n\times n$, es necesariamente una matriz escalar. (*Sugerencia: A* es conmutativa con cada E_{ij}.)

12. Si $\mathfrak{A}$ es un álgebra, demostrar que el conjunto $\mathfrak{J}$ de todos aquellos elementos z en $\mathfrak{A}$ que son conmutativos con todos los elementos de $\mathfrak{A}$ forman una subálgebra de $\mathfrak{A}$ (la cual es llamada el *centro* de $\mathfrak{A}$).

13. Determinar, a menos de isomorfismos, todas las álgebras lineales de orden dos sobre el campo real.

CAPÍTULO IX

Grupos lineales

1. Los grupos lineal y afín

Todas las transformaciones lineales no singulares en un espacio vectorial n-dimensional $V_n(F)$ forman un grupo, porque los productos y las inversas de tales transformaciones son también lineales y no singulares (Cap. VIII, Teor. 7). Este grupo será designado como *grupo lineal* (*) $L_n = L_n(F)$. En la correspondencia biunívoca entre transformaciones lineales y matrices, los productos se corresponden con productos, así que el grupo lineal $L_n(F)$ es isomorfo con el grupo de todas las matrices $n \times n$ regulares con elementos en el campo F.

Las traslaciones constituyen otro grupo importante. Una traslación del plano desplaza a todos los puntos de él en una misma distancia y según una misma dirección. La distancia y la dirección pueden ser representadas por un vector κ de magnitud y sentido apropiados. La traslación llevará entonces el extremo de cada vector ξ al extremo del vector $\xi+\kappa$. Pues bien, en un espacio $V_n(F)$ llamaremos *traslación* a la transformación $\xi \to \xi+\kappa$, con κ fijo. Con referencia a un sistema de coordenadas dado, las coordenadas y_i del vector transformado serán $y_1 = x_1 + k_1, \ldots, y_n = x_n + k_n$, donde las k_i son las coordenadas de κ. El producto de las dos traslaciones $\xi \to \eta = \xi+\kappa$, $\eta \to \zeta = \eta + \lambda$ es otra traslación, a saber, $\xi \to \zeta = \xi + (\kappa+\lambda)$; se corresponde, pues, exactamente con la suma de los vectores κ y λ.

(*) Se dirá *grupo lineal completo* cuando pueda haber confusión con alguno de sus subgrupos.

De modo semejante, la inversa de la traslación $\xi \to \xi + \kappa$ es $\eta \to \eta - \kappa$. Así hemos demostrado un caso particular del teorema de Cayley (Capítulo VI):

Teorema 1. *Todas las traslaciones $\xi \to \xi + \kappa$ de V_n forman un grupo abeliano isomorfo con el grupo aditivo de los vectores κ de V_n.*

Una transformación lineal T seguida de una traslación da

(1) $$\xi \to \eta = \xi T + \kappa \qquad (T \text{ lineal, } \kappa \text{ vector fijo}).$$

Se llamará transformación afín H de $V_n(F)$ a cualquier trans- 'ormación de esta forma. Las transformaciones afines (o afinidades) ıcluyen a las lineales (con $\kappa = 0$) y a las traslaciones (con $T = I$). Si .na transformación afín (1) es seguida de una segunda, $\eta \to \eta U + \lambda$, l producto será:

(2) $$\xi \to (\xi T + \kappa)U + \lambda = \xi(TU) + (\kappa U + \lambda).$$

El resultado es también afín, porque $\kappa U + \lambda$ es un vector de- erminado de $V_n(F)$. Cualquier traslación es biunívoca y tiene in- 'ersa, luego la transformación afín (1) será biunívoca si, y sólo si, .o es su parte lineal T. Su inversa será la transformación afín $\eta \to \xi = \eta T^{-1} - \kappa T^{-1}$, que se obtiene resolviendo (1) respecto a ξ. Esto demuestra:

Teorema 2. *Todas las transformaciones afines no singulares de $V_n(F)$ constituyen un grupo, que designaremos grupo afín $A_n(F)$. En él son subgrupos el grupo lineal y el grupo de traslaciones.*

¿Cuáles son, nos preguntamos ahora, las ecuaciones de una transformación afín referida a una base dada? La parte lineal T determina una matriz $A = \| a_{ij} \|$; las coordenadas del vector traslación constituyen una matriz de una fila $K = (k_1, \ldots, k_n)$. La afinidad transformará el vector de coordenadas $X = (x_1, \ldots, x_n)$ en un vector de coordenadas,

(3) $$Y = XA + K, \qquad y_j = \sum_{i=1}^{n} x_i a_{ij} + k_j.$$

Una transformación será afín si, y sólo si, se expresa con relación a alguna base, por ecuaciones lineales, aunque no homogéneas, tales como (3).

El producto de la transformación (3) por $Z=YB+L$ será:

(4) $\quad Z=X(AB)+KB+L \qquad (K,\ L,$ matrices de una fila),

fórmula que es correlativa de la (2). La misma regla de multiplicación se aplica a la matriz de orden $n+1$ construída a partir de (3) orlando la matriz A, por una columna de ceros y por la fila K seguida de un 1, es decir,

(5) $$\{Y=XA+K\} \leftrightarrow \begin{pmatrix} A & 0 \\ K & 1 \end{pmatrix} \qquad \begin{matrix} (0 \text{ es } n\times 1) \\ (K \text{ es } 1\times n) \end{matrix}$$

Porque, en efecto, la regla para multiplicar matrices vista en Capítulo VIII, § 3 (38), da

(6) $$\begin{pmatrix} A & 0 \\ K & 1 \end{pmatrix}\begin{pmatrix} B & 0 \\ L & 1 \end{pmatrix} = \begin{pmatrix} AB+0\cdot L & A\cdot 0+0\cdot 1 \\ KB+1\cdot L & K\cdot 0+1\cdot 1 \end{pmatrix} = \\ = \begin{pmatrix} AB & 0 \\ KB+L & 1 \end{pmatrix};$$

cuyo resultado es precisamente la matriz orlada que corresponde a la transformación producto (4). Las matrices regulares orladas forman un grupo (¿por qué?). Esto demuestra el siguiente

TEOREMA 3. *El grupo de las transformaciones afines no singulares de un espacio* n*-dimensional es isomorfo con el grupo de todas sus matrices* $(n+1)\times(n+1)$ *regulares cuya última columna es* $(0, \ldots, 0, 1)$. *El isomorfismo está dado explícitamente por la correspondencia (5).*

Cada transformación afín $\xi H=\xi T+\kappa$ determina una transformación lineal única T, y el producto de dos transformaciones afines determina, como en (2), el producto de sus correspondientes partes lineales. La correspondencia $H\to T$ representa el grupo de transformaciones afines regulares sobre el grupo lineal, y es un homomorfismo en el sentido de la teoría de grupos (Cap. VI, § 12). En todo homomorfismo, los elementos que tienen por homólogo a la identidad forman un subgrupo normal; pero en este caso, las afinidades H con $T=I$ son exactamente las traslaciones. Esto demuestra:

TEOREMA 4. *El grupo de las traslaciones es un subgrupo normal del grupo afín.*

EJERCICIOS

1. a) Representar cada una de las siguientes transformaciones afines por una matriz:

$$H_1: \quad x' = 3x + 6y + 2, \quad y' = 3y - 4;$$
$$H_2: \quad x' = x + y + 3, \quad y' = x - y + 5.$$

b) Calcular los productos H_1H_2, H_2H_1.
c) Hallar los inversos de H_1 y H_2.

2. a) Escribir explícitamente las ecuaciones de una transformación general afín del plano.
b) Sin necesidad de utilizar matrices, calcular, por sustitución, el producto de dos transformaciones afines cualesquiera del plano. Mostrar entonces que el resultado confirma la fórmula 4.
3. Dado el círculo $x^2 + y^2 = 1$, demostrar que cualquier transformación afín del plano transforma este círculo en una elipse o en un círculo.
4. Mostrar que el grupo de todas las traslaciones de $V_n(F)$ es isomorfo con un grupo de matrices.
5. a) Si J_2 es el campo de enteros mód. 2, hallar todas las matrices en $L_2(J_2)$.
b) Calcular la tabla de multiplicación para este grupo $L_2(J_2)$.

* 6. ¿Cuál es el orden del grupo lineal $L_2(F)$ cuando F es el campo de enteros módulo p?
7. Demostrar que la correspondencia $A \to (A')^{-1}$ es un automorfismo del grupo lineal.
8. Sea G el grupo de todas las matrices $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ con $ad \neq 0$. Mostrar que la correspondencia $A \to a$ es un homomorfismo.
9. Representar el grupo de las matrices 3×3 triangulares regulares (§ 3, siguiente) homomórficamente sobre las matrices 2×2 triangulares regulares. (*Sugerencia:* Proceder como en ejercicio 8, pero utilizando *bloques.*)
10. Si dos campos F y K son isomorfos, demostrar que los grupos $L_n(F)$ y $L_n(K)$ son isomorfos.
11. Si $n < m$, demostrar que $L_n(F)$ es isomorfo con un subgrupo de $L_m(F)$.
12. Demostrar que la identidad es la única transformación afín conmutable con cualquier otra transformación afín. (*Sugerencia:* Admitir el resultado del Ejerc. 11, Cap. VIII, § 6.)

*13. Si $L_n(F)$ es el grupo lineal completo, demostrar que dos transformaciones afines H_1 y H_2 están ambas en el mismo cogrupo a la derecha de $L_n(F)$ si, y sólo si, $OH_1 = OH_2$ (¡O es el origen!).

*14. Demostrar que el grupo cociente $A_n(F)/T_n(F)$ es isomorfo con $L_n(F)$, donde A_n indica el grupo afín y T_n el grupo de traslaciones.

2. Los grupos ortogonal y euclídeo

En la geometría de Euclides, la longitud desempeña un papel esencial. Vamos, por lo tanto, a fijarnos en aquellas transforma-

ciones de un espacio vectorial euclídeo que conservan las longitudes de todos los vectores.

DEFINICIÓN. *Una transformación lineal* T *se llama ortogonal si conserva el valor absoluto de cualquier vector* ξ, *así que* $|\xi T| = |\xi|$.

En el plano, esto sucede ciertamente en una rotación, o en una rotación seguida de una reflexión ; de acuerdo con la fórmula (2) del Cap. VIII, estas dos transformaciones están representadas, respectivamente, por las matrices

$$(7) \qquad \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}, \qquad \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}.$$

Estas dos matrices dan las únicas transformaciones ortogonales del plano real, según vamos a demostrar. Bajo cualquier transformación ortogonal $Y = XA$, los vectores unidad $\epsilon_1 = (1, 0)$ y $\epsilon_2 = (0, 1)$ tienen por transformados

$$(1, 0)\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = (a_1, a_2), \qquad (0, 1)\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = (b_1, b_2),$$

cuyas longitudes deben valer 1. De acuerdo con la fórmula pitagórica de la longitud, esto significa

$$(8) \qquad a_1^2 + a_2^2 = 1, \qquad b_1^2 + b_2^2 = 1.$$

Sumando, el vector (1, 1) tieine como transformado $(a_1 + b_1, a_2 + b_2)$ de longitud $\sqrt{2}$, luego $(a_1 + b_1)^2 + (a_2 + b_2)^2 = 2$. Desarrollando, y en virtud de (8), nos resultará

$$(9) \qquad a_1 b_1 + a_2 b_2 = 0.$$

En virtud de (8), hay un ángulo θ con $\cos\theta = a_1$, $\operatorname{sen}\theta = a_2$. Entonces es $\operatorname{tg}\theta = a_2/a_1 = -b_1/b_2$, por (9), mientras que por (8) es $b_2 = \pm\cos\theta$, $b_1 = \mp\operatorname{sen}\theta$. Las dos posibilidades de signo dan exactamente las dos matrices (7).

La inversa de una rotación de amplitud θ es la rotación de amplitud $-\theta$, así que la matriz inversa es

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos(-\theta) & \operatorname{sen}(-\theta) \\ -\operatorname{sen}(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

¡La matriz que resulta es precisamente la transpuesta de la original! Esta circunstancia será generalizada a las matrices $n \times n$ ortogonales.

TEOREMA 5. *Una transformación ortogonal T tiene, para cualquier par de vectores, ξ, η, las propiedades siguientes:*

1) T *conserva las distancias, ó* $|\xi - \eta| = |\xi T - \eta T|$.
2) T *conserva los productos internos, ó* $(\xi, \eta) = (\xi T, \eta T)$.
3) T *conserva la ortogonalidad, ó* $\xi \perp \eta$ *implica* $\xi T \perp \eta T$.
4) T *conserva la magnitud angular, ó* $\cos \angle(\xi, \eta) = \cos \angle(\xi T, \eta T)$

Demostración. Como T es lineal, la definición da 1). Como $\xi \perp \eta$ significa $(\xi, \eta) = 0$ y puesto que los ángulos son expresables mediante productos internos [Cap. VII, (28)] las propiedades 3) y 4) se deducen inmediatamente de 2). Pero para 2), de la bilinealidad del producto interno resulta que $(\xi + \eta, \xi + \eta) = (\xi, \xi) + 2(\xi, \eta) + (\eta, \eta)$. Esta ecuación permite expresar (ξ, η) mediante «longitudes», pues siendo $|\xi| = (\xi, \xi)^{1/2}$, resulta

$$2(\xi, \eta) = |\xi + \eta|^2 - |\xi|^2 - |\eta|^2. \tag{10}$$

Ahora bien: la transformación ortogonal T deja invariantes las longitudes del segundo miembro de esta igualdad, luego también el producto interno del primer miembro, lo cual demuestra 2).

Recíprocamente, una transformación T que conserva todos los productos internos, debe conservar las longitudes, y por tanto ser ortogonal, ya que la longitud es expresable por un producto interno.

Investiguemos ahora qué matrices corresponden a las transformaciones ortogonales lineales. Este problema es fácil, al menos con referencia a una base ortogonal normal.

TEOREMA 6. *Con relación a cualquier base ortogonal normal, una matriz* A *representa una transformación lineal ortogonal si, y sólo si, considerando las filas de* A *como vectores, cada fila tiene longitud unidad, y dos filas cualesquiera son ortogonales.*

Demostración. Cualquier transformación ortogonal debe, por el Teorema 5, transformar la base dada $\epsilon_1, \ldots, \epsilon_n$, en una nueva base, $\alpha_1 = \epsilon_1 T, \ldots, \alpha_n = \epsilon_n T$, también normal y ortogonal. Recíprocamente, supongamos que T tiene tal propiedad y sea $\xi = x_1\epsilon_1 + \ldots + x_n\epsilon_n$ un

vector cualquiera ; su transformado será $\xi T = x_1\alpha_1 + \ldots + x_n\alpha_n$; sabemos, por el Teorema 15 del Cap. VII, que la longitud está dada por la fórmula

$$|\xi| = (x_1^2 + \ldots + x_n^2)^{1/2} = |\xi T|,$$

luego T es ortogonal. La demostración se completa con la observación (cfr. Cap. VIII, § 4) de que la fila i-ésima de A, considerada como un vector, representa las coordenadas de $\alpha_i = \epsilon_i T$ relativas a la base original $\epsilon_1, \ldots, \epsilon_n$.

DEFINICIÓN. *Una matriz cuadrada* A *se llama ortogonal si, y sólo si, cada fila de* A *tiene longitud 1, y dos filas cualesquiera son ortogonales. En fórmula, esto significa que* $A = \| a_{ij} \|$ *es ortogonal si, y sólo si,*

$$(11) \qquad \sum_{k=1}^{n} a_{ik}a_{ik} = 1 \text{ para todo } i, \qquad \sum_{k=1}^{n} a_{ik}a_{jk} = 0 \text{ si } i \neq j.$$

Las condiciones (11) son exactamente las mismas que han sido desarrolladas en (8) y (9) para matrices cuadradas 2×2. Si representamos por A_i la fila i-ésima de la matriz A, y por A_i' su transpuesta, el producto interior de A_i por A_i' es la matriz producto A_iA_i' (ver (30), Cap. VIII), así que las condiciones (11) pueden escribirse como sigue :

$$(11') \qquad A_iA_i' = 1, \quad A_iA_j' = 0 \quad (i \neq j).$$

Designando A' la transpuesta de A, las fórmulas (11') nos dicen que la matriz producto AA' (que se calcula multiplicando filas por columnas) tiene en el cruce de la fila i con la columna j el elemento $A_iA_j' = \delta_{ij}$ que ocupa el mismo lugar en la matriz idéntica I cuyos elementos diagonales son $\delta_{11} = \delta_{22} = \ldots = \delta_{nn} = 1$, siendo nulos los restantes. Hemos demostrado, pues,

TEOREMA 7. *Una matriz* A *es ortogonal si, y sólo si,* $AA' = I$.

Este resultado significa que la transpuesta A' de una matriz ortogonal A es inversa por la derecha de A, luego por el Teor. 7 del Cap. VIII, cualquier matriz ortogonal A es regular, con $A^{-1} = A'$. Por lo tanto, $A'A = I$. Esta igualdad puede expresarse diciendo que $A'(A')' = I$ cuando A' es ortogonal ; luego la transpuesta de cualquier matriz ortogonal A es también ortogonal. De esto se sigue que

una matriz A es ortogonal si, y sólo si, cada columna de A tiene longitud 1, y dos columnas cualesquiera son ortogonales,

$$(12) \qquad \sum_{k=1}^{n} a_{ki}a_{ki} = 1 \text{ para todo } i, \qquad \sum_{k=1}^{n} a_{ki}a_{kj} = 0 \text{ si } i \neq j.$$

Todas las matrices $n \times n$ ortogonales forman grupo, pues la inversa A^{-1} de una matriz ortogonal lo es asimismo $[(A^{-1})' = (A')^{-1}]$, y el producto de dos matrices ortogonales A y B es ortogonal $[(AB)' = B'A' = B^{-1}A^{-1} = (AB)^{-1}]$. Este subgrupo del grupo lineal $L_n(R^*)$ se llama el *grupo ortogonal* O_n; es isomorfo con el grupo de todas las transformaciones ortogonales del espacio euclídeo considerado.

Se llama *movimiento rígido* de un espacio vectorial euclídeo E, una transformación no singular U de E que conserve las distancias, esto es, que satisfaga a la igualdad $|\xi U - \eta U| = |\xi - \eta|$ para todos vectores ξ y η. Cualquier traslación de E conserva los vectores diferencia $|\xi - \eta|$, luego también su longitud, y, por lo tanto, es un movimiento rígido. Por consiguiente, si una transformación afín $\xi \to \xi T + \kappa$ es rígida, también lo es $\xi \to (\eta - \kappa) = \xi T$; recíprocamente, si T es rígida, también lo es $\xi \to \eta = \xi T + \kappa$. Pero, por el Teorema 5, una transformación lineal es rígida si, y sólo si, es ortogonal. En conclusión, *una transformación afín (1) es un movimiento rígido si, y sólo si,* T *es ortogonal.* Sigue de aquí, como en la demostración del Teorema 2, ya que todas las transformaciones ortogonales forman un grupo, que la totalidad de las transformaciones afines rígidas constituye un subgrupo del grupo afín, llamado *grupo euclídeo.* Este constituye la base de la geometría euclídea (*).

Aparecen en geometría otros grupos. Es muy conocido, por ejemplo, el grupo de las transformaciones por semejanza, constituído por aquellas transformaciones lineales T que alteran las longitudes multiplicándolas por un factor numérico constante $c_T \neq 0$, de modo que $|\xi T| = c_T |\xi|$. Es fácil probar que efectivamente constituyen un grupo, del cual es un subgrupo el grupo ortogonal. Más amplio aún es el grupo equiforme, constituído por las transposiciones afines $\xi \to \xi T + \kappa$ en donde T representa una transformación por semejanza.

(*) Como cualquier transformación rígida es necesariamente afín, resulta que el grupo euclídeo es el grupo de todos los movimientos rígidos.

EJERCICIOS

1. Comprobar la ortogonalidad de las siguientes matrices. Si una de ellas es ortogonal, hallar su inversa.

a) $\begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix}$, b) $\begin{pmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$, c) $\begin{pmatrix} 0{,}6 & 0{,}8 \\ 0{,}8 & -0{,}6 \end{pmatrix}$

2. Hallar una matriz ortogonal con la primera fila múltiplo escalar de (5, 12, 0).
3. Demostrar que, si las columnas de una matriz ortogonal se permutan, la que resulta también es ortogonal.
4. Si A y B son ortogonales, demostrar que $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ lo es también.
5. Multiplicar las dos matrices que siguen y comprobar la ortogonalidad del producto.

$$\begin{pmatrix} \cos\phi & \operatorname{sen}\phi & 0 \\ -\operatorname{sen}\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta \\ 0 & -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

6. Demostrar que el grupo euclídeo es isomorfo con un grupo de matrices.
7. Demostrar que todas las traslaciones constituyen un subgrupo normal del grupo euclídeo.
8. Como demostración de 2) en Teorema 5, demostrar a partir de los principios que $4(\xi, \eta) = |\xi+\eta|^2 - |\xi-\eta|^2$.
9. Demostrar geométricamente que cualquier transformación ortogonal del plano es una rotación o una reflexión (simetría respecto a un eje) seguida por una rotación. (*Sugerencia:* Una transformación lineal está completamente determinada por las imágenes de dos vectores perpendiculares unitarios.)
10. Demostrar que cualquier transformación de semejanza S puede escribirse en la forma $S=cT$ como producto por un escalar positivo de una transformación ortogonal T, y esto de un solo modo.
11. Dar condiciones necesarias y suficientes para que la matriz A represente una transformación de semejanza referida a una base ortogonal y normal (cfr. Teors. 6-7).
12. a) Demostrar que todas las transformaciones de semejanza constituyen un grupo S_n.
    b) Demostrar que O_n es un subgrupo normal de S_n.
    c) Demostrar que el grupo cociente S_n/O_n es isomorfo con el grupo multiplicativo de los números reales positivos.
13. a) Demostrar que las matrices $n\times n$ que satisfacen a $AA'=I$ con coeficientes en cualquier campo, forman un grupo multiplicativo.
    b) ¿Cuántas matrices 3×3 «ortogonales» existen, con coeficientes en J_3?

3. Matrices diagonales y de permutación

Entre los muchos grupos de matrices que es posible considerar, aparecen útilmente aquellos en que intervienen matrices de forma

«diagonal» o de otras formas sencillas. Una matriz $D = \| d_{ij} \|$ se llama *diagonal* cuando $i \neq j$ implica $d_{ij}=0$, esto es, cuando los elementos no nulos de D estén situados en la diagonal principal (desde el ángulo superior izquierda al inferior derecha). Para sumar o multiplicar matrices diagonales, basta con sumar o multiplicar los elementos correspondientes alineados en tales diagonales (¿por qué?). De aquí resulta que una matriz diagonal tiene una inversa si ningún elemento de la diagonal es nulo, y sólo en este caso. Por lo tanto, resulta :

TEOREMA 8. *Todas las matrices diagonales regulares forman un subgrupo del grupo lineal.*

Una *matriz de permutación* P es una matriz que tiene en cada fila y en cada columna un solo elemento igual a la unidad y todos los restantes elementos son nulos.

Las matrices 3×3 de permutación son seis. Una de ellas es I, y las restantes :

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

En dos filas cualesquiera, los elementos 1 están en distinta columna, luego las filas son ortogonales ; se deduce, pues, que toda matriz de permutación P es ortogonal, y $P^{-1}=P'$. P es una matriz de permutación si, y sólo si, la transformación correspondiente $Y=XP$ efectúa una permutación de las unidades vectoriales $I_1, \ldots, I_n$. Por lo tanto, las matrices de permutación se corresponden de manera biunívoca con las $n!$ permutaciones de n símbolos (Capítulo VI, § 7), y esta correspondencia es un isomorfismo. Hemos, pues, demostrado

TEOREMA 9. *Las matrices de permutación* $n \times n$ *forman un subgrupo del grupo ortogonal, que es isomorfo con el grupo simétrico de* n *letras.*

Veamos ahora otras importantes clases de matrices. Una matriz M se llama *monomial* si cada fila y cada columna tiene precisamente un elemento distinto de cero ; así que estas matrices pueden

obtenerse sustituyendo los elementos 1 de una matriz de permutación por otros elementos no nulos, como, por ejemplo, en

$$(13)\quad M_1=\begin{pmatrix} 0 & 0 & 5\\ -2 & 0 & 0\\ 0 & 3 & 0\end{pmatrix},\quad M_2=\begin{pmatrix} 0 & 7 & 0\\ 0 & 0 & -3\\ 4 & 0 & 0\end{pmatrix},\quad M_3=\begin{pmatrix} 0 & 4\\ -1 & 0\end{pmatrix}.$$

Una matriz cuadrada $T=\|t_{ij}\|$ se llama triangular si todos los elementos debajo de la diagonal principal son nulos ; es decir, si $t_{ij}=0$ siempre que $i>j$. Una matriz S se dice *estrictamente triangular* cuando todos los elementos *en* o debajo de la diagonal principal son nulos. En el caso 4×4, estos dos tipos son, pues, los siguientes :

$$(14)\qquad T=\begin{bmatrix} q & r & s & t\\ 0 & u & v & w\\ 0 & 0 & x & y\\ 0 & 0 & 0 & z\end{bmatrix},\qquad S=\begin{bmatrix} 0 & u & v & w\\ 0 & 0 & x & y\\ 0 & 0 & 0 & z\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

donde las letras indican elementos arbitrarios. Finalmente, llamaremos *matriz escalar* la que puede escribirse en la forma cI, siendo I la matriz idéntica.

La distribución según cierto «patrón» de los elementos no nulos de una matriz, no es el único medio de construir grupos de matrices. Cualquier grupo de transformaciones lineales puede ser representado por el correspondiente grupo de matrices. Por ejemplo, el grupo del cuadrado consta de transformaciones lineales. Elijamos como origen el centro del cuadrado, y tomemos los ejes paralelos a sus lados. Si formulamos mediante las coordenadas $(x,\ y)$ los movimientos R, R', H y D (ver su descripción en Cap. VI, § 1), obtenemos sendas transformaciones con las siguientes matrices :

$$R\to\begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix},\quad R'\to\begin{pmatrix} -1 & 0\\ 0 & -1\end{pmatrix},\quad H\to\begin{pmatrix} 1 & 0\\ 0 & -1\end{pmatrix},\quad D\to\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}.$$

y los otros cuatro elementos del grupo pueden representarse análogamente. La tabla de multiplicación de este grupo, dada en Cap. VI, § 4, puede ser calculada por simple multiplicación de las matrices correspondientes (¡ compruébese !). En otras palabras, el grupo del cuadrado es isomorfo con un grupo de ocho matrices 2×2.

EJERCICIOS

(Algunos ejercicios sobre matrices diagonales aparecen en Cap. VIII, § 2.)

1. a) Mostrar dos matrices 4×4 de permutación.
   b) Exhibir tres matrices monomiales que no sean ni diagonales ni matrices de permutación.
2. Mostrar explícitamente el isomorfismo entre las matrices de permutación 3×3 y el grupo simétrico.
3. Sea S_i el subespacio unidimensional de V_n engendrado por el i-ésimo vector básico ϵ_i. Demostrar que una matriz regular D es diagonal si, y sólo si, la transformación lineal correspondiente $Y = XD$ representa a cada espacio S_i sobre sí mismo.
4. Encontrar para las matrices monomiales una caracterización análoga a la del Ejerc. 3.
5. a) Demostrar que una matriz monomial M puede escribirse de modo único en la forma $M = DP$, donde D es regular y diagonal y P es matriz de permutación. (*Sugerencia:* Recordar Ejerc. 11, § 2, Cap. VIII.)
   b) Escribir las matrices M_1 y M_2 del texto en las formas DP y PD.
6. Describir la inversa de una matriz monomial M y hallar las inversas de M_1 y M_2 en (13).
7. Si M es monomial y D diagonal, demostrar que $M^{-1}DM$ es diagonal.
8. Si P es una matriz de permutación y D diagonal, describir la forma de la transformada $P^{-1}DP$.
9. ¿Cuántas filas tiene PA, según las de A?
10. Una matriz A se llama *nilpotente* si alguna potencia de A es 0. Demostrar que cualquier matriz estrictamente triangular es nilpotente. (*Sugerencia:* Probar con el caso 3×3.)
11. ¿Cuáles de los siguientes conjuntos de matrices son grupos multiplicativos?
    a) Todas las matrices escalares cI;
    b) Todas las matrices diagonales;
    c) Todas las matrices diagonales regulares;
    d) Todas las matrices de permutación;
    e) Todas las matrices monomiales;
    f) Todas las matrices triangulares;
    g) Todas las matrices estrictamente triangulares;
    h) Todas las matrices con ceros en la segunda fila;
    i) Todas las matrices en las que al menos una fila consiste en ceros.

*12. ¿Cuáles de los conjuntos de matrices del Ejerc. 11 son subálgebras del álgebra de las matrices $M_n(F)$? Si un conjunto es un subálgebra, determinar su orden.

13. Representar el grupo de las simetrías del rectángulo como un grupo de matrices.
14. Hallar las matrices que representan todas las simetrías del cuadrado, calculando, por el producto de matrices, los productos HR, RH, HD y DD', y comparar estos resultados con los de la tabla de multiplicación del § 4 del Cap. VI.

15. a) Demostrar que todas las transformaciones biunívocas $y=(ax+b)/(cx+d)$ con $ad\neq bc$ forman un grupo (llamado *grupo homográfico*).
b) Demostrar que este grupo es isomorfo con el grupo cociente del grupo lineal módulo el subgrupo de las matrices escalares no nulas.
* c) Extender este resultado a matrices de tipo superior al 2×2.

16. a) Demostrar que el conjunto de todas las matrices regulares de la forma $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ con A $r\times r$ y B $s\times s$, es un grupo (r y s son enteros fijos).
* b) ¿Cuál es el carácter geométrico de las transformaciones lineales de $V_3(R^3)$ determinadas por una matriz de éstas, si $r=2$, $s=1$?

4. Cambio de base

En el plano, las ecuaciones

$$(15) \qquad y_1=x_1+a, \qquad y_2=x_2+b$$

pueden interpretarse de dos modos: como una transformación que lleve el punto de coordenadas (x_1, x_2) a un nuevo punto de coordenadas (y_1, y_2); o como un cambio de coordenadas, en el cual el punto P de coordenadas (x_1, x_2) con referencia a unos ejes de origen O, está determinado por otras coordenadas (y_1, y_2) con relación a otros ejes paralelos con un nuevo origen. [Las coordenadas del nuevo origen, en el sistema primitivo son $(-a, -b)$, porque (15) dará para coordenadas nuevas de este punto $(-a+a, -b+b) = (0, 0)$.]

Estas dos interpretaciones pueden llamarse *activa* (el punto se mueve) y *pasiva* (el punto permanece quieto). Un doble significado similar va envuelto en las restantes ecuaciones de transformación.

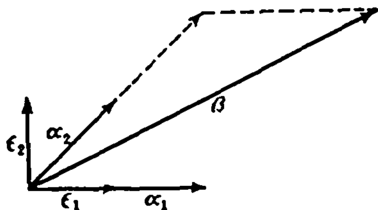


Figura 1

Cualquier cambio de base implica un cambio de coordenadas. Por definición, las coordenadas de un vector con relación a una base son los coeficientes de la expresión lineal del vector en función de los de la base. Por lo tanto, dos bases $\epsilon_1, \ldots, \epsilon_n$ y $\alpha_1, \ldots, \alpha_n$ de un espacio vectorial dan dos conjuntos de coordenadas x_i y x_i^* para un mismo vector ξ,

$$(16) \qquad \xi=x_1\epsilon_1+\ldots+x_n\epsilon_n=x_1^*\alpha_1+\ldots+x_n^*\alpha_n.$$

Consideremos, por ejemplo, el vector $\beta=4\epsilon_1+2\epsilon_2$ en el plano, de coordenadas (4, 2) respecto a la base ϵ_1, ϵ_2. Los vectores

(17) $$\alpha_1=2\epsilon_1, \qquad \alpha_2=\epsilon_1+\epsilon_2$$

forman también una base, y β puede expresarse así: $\beta=\alpha_1+2\alpha_2$, de modo que las nuevas coordenadas, referidas a estos ejes «oblicuos» (ver fig. 1), son 1 y 2. Para cualquier vector $\xi=x_1^*\alpha_1+x_2^*\alpha_2$ con nuevas coordenadas x_1^*, x_2^* por la sustitución (17) resultará:

$$\xi=x_1^*(2\epsilon_1)+x_2^*(\epsilon_1+\epsilon_2)=(2x_1^*+x_2^*)\epsilon_1+x_2^*\epsilon_2.$$

Los coeficientes de ϵ_1 y de ϵ_2 son, por definición, las coordenadas antiguas x_i de ξ, así que están dadas por las igualdades

(18) $$x_1=2x_1^*+x_2^* \qquad x_2=x_2^*$$

Las nuevas coordenadas pueden expresarse mediante las antiguas resolviendo estas ecuaciones, y resulta $x_2^*=x_2$, $x_1^*=(x_1-x_2)/2$. Estas expresiones para las nuevas coordenadas son exactamente análogas a las ecuaciones de una transformación lineal.

Consideremos más generalmente una nueva base, expresada en función de la antigua mediante las ecuaciones

(19) $$\alpha_i=a_{i1}\epsilon_1+\ldots+a_{in}\epsilon_n=\sum_j a_{ij}\epsilon_j \qquad (i=1, \ldots, n).$$

La fila i de la matriz de coeficientes $A=\|a_{ij}\|$ es la fila de las coordenadas antiguas $(a_{i1}, \ldots, a_{in})$ del vector α_i. Como estos vectores forman una base, las filas son independientes, y la matriz A es regular (Cap. VIII, Teor. 7). Sustituyendo (19) en la nueva expresión $\xi=\Sigma x_i^*\alpha_i$ de un vector ξ, tendremos

$$\xi=\sum_i x_i^*\left(\sum_j a_{ij}\epsilon_j\right)=\sum_j\left(\sum_i x_i^*a_{ij}\right)\epsilon_j.$$

Los coeficientes de las ϵ_j son las antiguas coordenadas

(20) $$x_j=x_1^*a_{1j}+\ldots+x_n^*a_{nj}=\sum_{i=1}^n x_i^*a_{ij}, \qquad j=1, \ldots, n.$$

Estas ecuaciones pueden escribirse en forma matricial $X=X^*A$ o $X^*=XA^{-1}$, exactamente como en las ecuaciones de una transformación lineal biunívoca.

La relación entre las ecuaciones para las «bases» (19) y para las «coordenadas» (20), puede establecerse de modo conveniente

utilizando las matrices de una fila $\alpha=(\alpha_1, \ldots, \alpha_n)$ y $\epsilon=(\epsilon_1, \ldots, \epsilon_n)$. Entonces, en (19) aparece la multiplicación de A por la transpuesta ϵ' (matriz de una columna), o sea, $\alpha'=A\epsilon'$. Por transposición, esta fórmula de $\alpha=\epsilon A'$. Las nuevas bases y coordenadas son, pues, expresadas mediante las antiguas como sigue:

(21) Bases: $\alpha=\epsilon A'$, Coordenadas: $X^*=XA^{-1}$.

La matriz A^{-1} de la segunda ecuación es la inversa de la transpuesta de la matriz A' de la primera. Estos hechos se expresan diciendo que la transformación de coordenadas es *contravariante* de la transformación de las bases. Cualquier matriz regular A proporcionará una posible base nueva para el espacio, $\alpha=\epsilon A'$, y asimismo cualquier ecuación $X^*=XB$, con $B=A^{-1}$ regular, representará un posible grupo de nuevas coordenadas X^*.

Podemos ahora establecer la doble interpretación de una ecuación matricial $Y=XB$, con B regular. Como *activa*, esta ecuación representa una transformación lineal no singular T, que transporta cada vector ξ con coordenadas X relativas a la base ϵ, sobre un vector η de coordenadas $Y=XB$. Como *pasiva*, esta ecuación da las nuevas coordenadas Y de un vector ξ con coordenadas primitivas X referidas a la base ϵ; estas nuevas coordenadas se toman con relación a una nueva base α expresada por $\alpha=\epsilon A'$, siendo A la inversa de la matriz B. Este resultado se puede enunciar así:

TEOREMA 10. *Cualquier sustitución no singular* $Y=XB$ *sobre* n *variables* $x_1, \ldots, x_n$, *puede ser interpretada como una transformación lineal biunívoca o como un cambio de coordenadas, e inversamente. En consecuencia, toda simplificación algebraica de una expresión dada, que puede conseguirse por una conveniente transformación del espacio* $V_n(F)$, *puede asimismo ser conseguida por un conveniente cambio de coordenadas en tal espacio, y recíprocamente.*

En un espacio vectorial euclídeo, la nueva base $\alpha_1, \ldots, \alpha_n$ de (19), será una base ortogonal y normal si, y sólo si, las filas de coordenadas $(a_{i1}, \ldots, a_{in})$ son mutuamente ortogonales y cada una de longitud unidad, según las fórmulas usuales. Esto significa que la matriz A de los coeficientes es ortogonal (Teorema 6); en tal caso, también la inversa A^{-1} es ortogonal e igual a A'.

COROLARIO. *Un cambio de coordenadas, desde una antigua a una nueva base ortogonal normal de un espacio euclídeo, corresponde como en (21) a una transformación ortogonal del espacio, y recíprocamente. Por lo tanto, una transformación ortogonal y la elección de un nuevo sistema ortogonal y normal de coordenadas, tienen los mismos efectos, en el sentido del Teorema 10.*

Esta conclusión será aplicada en la próxima sección para la simplificación de funciones lineales.

Para la geometría afín, los resultados son semejantes. Un nuevo sistema de coordenadas se establecerá eligiendo como origen otro punto P_0 (que será extremo de cierto vector λ) y como nueva base otros vectores independientes $\alpha_1, \dots, \alpha_n$. Las coordenadas de cualquier ξ relativas a este nuevo sistema son los números y_i dados por $\xi - \lambda = \Sigma y_i \alpha_i$; dicho de otro modo, empleamos las componentes del vector $\xi - \lambda$ que une al nuevo origen con el extremo de ξ. Las ecuaciones que expresan estas nuevas coordenadas en función de las antiguas, corresponden a una transformación afín biunívoca, y recíprocamente, las ecuaciones de una tal transformación pueden ser interpretadas como ecuaciones de un conveniente cambio afín de coordenadas.

EJERCICIOS

1. Si los vectores indicados se toman como una nueva base vectorial del espacio, hallar las ecuaciones que corresponden a las coordenadas antiguas en función de las nuevas. En los casos a) y b) dibujar una figura.

   a) $\alpha_1 = (1, 1)$, $\alpha_2 = (1, -1)$;

   b) $\alpha_1 = (2, 3)$, $\alpha_2 = (-2, -1)$;

   c) $\alpha_1 = (1, 1, 0)$, $\alpha_2 = (1, 0, 1)$, $\alpha_3 = (0, 1, 1)$.

2. Si una nueva base β_j se da indirectamente mediante ecuaciones de la forma $\epsilon_i = \Sigma_j b_{ij} \beta_j$, investigar las ecuaciones del correspondiente cambio de coordenadas.
3. Presentar las ecuaciones para el cambio de coordenadas, debido a una rotación de los ejes en el plano.
4. Discutir la traslación de coordenadas en el espacio afín.
5. Dar una exposición detallada de las correspondencias biunívocas entre los cambios de base afines y las transformaciones afines.
6. a) Demostrar que la correspondencia $A \to \theta(A) = (A^{-1})'$ es un automorfismo del grupo lineal $L_n(F)$.

   b) Demostrar que $\theta^2(A) = A$ para todo A.

   c) En el caso $F = R^*$, ¿para qué matrices es $\theta(A) = A$?

5. Equivalencia y formas canónicas. Invariantes

Una importante aplicación de los grupos lineales es su empleo en la simplificación de polinomios lineales y cuadráticos. La idea citada se puede ilustrar con el conocido proceso de «completar el cuadrado» en el trinomio $f(x)=ax^2+bx+c$, con $a\neq 0$. Este proceso enseña que la sustitución $y=x+b/2a$ transforma $f(x)$ en el nuevo polinomio

(22) $$g(y)=ay^2-d/4a, \quad \text{con} \quad d=b^2-4ac.$$

Ante tal circunstancia, diremos que $f(x)$ *equivale* a $g(y)$ mediante la traslación (*) $y=x+b/2a$ (o que $f(x)$ y $g(y)$ son equivalentes para el grupo de las traslaciones). En general, si aplicamos cualquier traslación $z=x+k$ a $f(x)=ax^2+bx+c$, obtenemos

$$h(z)=a(z-k)^2+b(z-k)+c=az^2+(b-2ak)z+(ak^2-bk+c).$$

En el polinomio $h(z)$ que resulta, el término lineal puede eliminarse solamente cuando $k=b/2a$, como al completar el cuadrado. Por lo tanto, el polinomio $f(x)$ es equivalente, para el grupo de las traslaciones, a un solo polinomio ay^2+h sin término lineal. Por esta razón diremos que ay^2+h es una *forma canónica* respecto a las traslaciones.

Interesa también investigar los «invariantes» para esta equivalencia. Así, el *discriminante* d de $f(x)$ está dado por $d=b^2-4ac$. Si calculamos el discriminante d' del polinomio transformado $h(z)$ con los coeficientes $b'=b-2ak$, $c'=ak^2-bk+c$, se obtiene

$$d'=b'^2-4a'c'=b^2-4akb+4a^2k^2-4a^2k^2+4akb-4ac=b^2-4ac.$$

Por lo tanto, el discriminante es inalterable o *invariante* en las traslaciones. En función del discriminante, la forma canónica $g(y)$ puede escribirse $ay^2-d/4a$.

Para definir en general la noción de equivalencia respecto a un grupo, consideremos dos polinomios $f(x_1, \ldots, x_n)$ y $g(y_1, \ldots, y_n)$, cada uno de n indeterminadas, sobre un campo F. diremos que f es equivalente a g en un grupo dado G de transformaciones $X\to Y$

(*) Como $2=1+1$, este resultado es válido bajo la hipótesis de que en el campo de los coeficientes de $f(x)$ es $1+1\neq 0$. Ahora bien, el campo J_2 de los enteros módulo 2, muestra que existen campos en los que $1+1=0$. Estos campos (que se llaman de característica 2) son excepcionales en muchas partes de la teoría de matrices y determinantes.

del espacio n-dimensional $V_n(F)$ si, y sólo si, existe en el grupo una transformación $X \to Y$, que a la expresión de f le haga corresponder la expresión de g. Por ejemplo, f y g serán equivalentes en el grupo lineal si, y sólo si, f resulta igual a g por efecto de una sustitución $x_i = \Sigma y_j a_{ji}$, donde $A = \| a_{ji} \|$ es una matriz regular.

Entre todos los polinomios de un tipo determinado, es posible seleccionar algunos polinomios más particulares y tales, que cualquier polinomio del tipo en cuestión sea equivalente en un grupo G a uno y sólo a uno de los polinomios seleccionados. Estos últimos constituyen las formas canónicas del tipo dado para el grupo G. Es conveniente elegir las formas canónicas lo más sencillas que sea posible.

También puede definirse de modo general la noción de «invariante». Para cualquier polinomio f de un tipo dado, dispongamos en un orden determinado $b_1, \ldots, b_m$ los coeficientes. Entonces una función $J(f)$ expresada en función de estos coeficientes, $J(f) = J(b_1, \ldots, b_m)$, se llamará invariante de f en un grupo G, si $J(f) = J(g)$, siempre que f y g sean equivalentes en el grupo. Se dice que los invariantes $J_1, \ldots, J_r$ constituyen un sistema completo de invariantes de un cierto tipo de polinomios, cuando para la equivalencia de dos polinomios f y g de este tipo, es necesaria y suficiente la igualdad entre tales invariantes, así que $J_1(f) = J_1(g), \ldots, J_r(f) = J_r(g)$. Por ejemplo, en el caso de polinomios cuadráticos, para el grupo de las traslaciones, el primer coeficiente a y el discriminante d forman un sistema completo de invariantes, pues cualquiera de tales polinomios es equivalente a la forma canónica $ay^2 - d/4a$, según (22).

Para el grupo afín, consideremos el caso de los polinomios de forma lineal,

$$f(x_1, \ldots, x_n) = b_1 x_1 + \ldots + b_n x_n + c, \tag{23}$$

con coeficientes b_i y c de un campo F. Excluiremos el caso trivial de una constante suponiendo que algún coeficiente b_j es distinto de cero. La traslación $x_j' = x_j + c/b_j$, con $x_i' = x_i$ si $i \neq j$, hará desaparecer el término constante c. Además, la transposición $x_1'' = x_j'$, $x_j'' = x_1'$ nos da una nueva forma en la que el coeficiente b_j de x_1'' no es cero. Por lo tanto, podemos suponer en (23) que $b_1 \neq 0$ y $c = 0$. Ahora bien, la transformación dada por las ecuaciones

$$y_1 = b_1 x_1 + \ldots + b_n x_n + c, \qquad y_2 = x_2, \ \ldots, \ y_n = x_n \tag{24}$$

reemplazará f por la función equivalente $g(y_1, \ldots, y_n)=y_1$. La transformación expresada por (24) es afín, y es biunívoca por ser $b_1 \neq 0$. Por lo tanto, todas las formas lineales (23) son equivalentes, en el grupo afín, a la forma canónica $g(y_1, \ldots, y_n)=y_1$. Además, todas las formas (23) lineales *homogéneas* (con $c=0$) son equivalentes para el grupo lineal a la misma forma canónica.

Sobre cualquier campo F, distintas *formas* lineales (23) determinan distintas *funciones* lineales (cfr. Cap. IV, § 2); los coeficientes de la forma vienen determinados por los valores de la función como $f(1, 0, \ldots, 0)=b_1+c, \ldots, f(0, \ldots, 0, 1)=b_n+c, f(0, \ldots, 0)=c$. De igual modo, cuando $c=0$ la (23) es una función homogénea lineal $f(\xi)$ del vector ξ de coordenadas $x_1, \ldots, x_n$ relativas a la base $\epsilon_1, \ldots, \epsilon_n$ de $V_n(F)$. Por el Teorema 10, el precedente resultado sobre formas canónicas respecto al grupo lineal demuestra que, dada una función lineal homogénea $f(\xi)$, se puede elegir una nueva base $\alpha_1, \ldots, \alpha_n$ tal, que la expresión de f mediante las nuevas coordenadas sea $f(\xi)=x_1^*$. Esto significa que el conjunto S de todos los vectores ξ que satisfacen a la ecuación $f(\xi)=0$ es, simplemente, el conjunto de todos los vectores cuya primera coordenada $x_1^*=0$, de modo que S constituye un subespacio $(n-1)$ dimensional, engendrado por los vectores $\alpha_2, \ldots, \alpha_n$ de la nueva base. Así hemos demostrado que *si una función lineal homogénea* $\Sigma b_i x_i$ *de* n *variables no es idénticamente nula, sus ceros forman un subespacio* (n — 1) *dimensional (llamado hiperplano) en* $V_n(F)$.

Para considerar bajo el grupo euclídeo las funciones lineales de un espacio euclídeo, es lo más conveniente adoptar la interpretación *pasiva*. Como antes, con una traslación previa haremos $c=0$ en (23). La función lineal homogénea que resulta, $f(\xi)$, tiene estas dos propiedades características:

$$(25) \qquad f(\xi_1+\xi_2)=f(\xi_1)+f(\xi_2), \qquad f(d\xi)=df(\xi),$$

para dos vectores ξ_1, ξ_2 cualesquiera y cualquier escalar d. El hiperplano $f(\xi)=0$ posee una base ortogonal normal $\alpha_2, \ldots, \alpha_n$ que puede ampliarse (Cap. VII, § 9, Lema 2) a una base ortogonal normal $\alpha_1, \alpha_2, \ldots, \alpha_n$ de la totalidad del espacio E. Con relación a esta base, cualquier vector ξ puede expresarse como en (16), con coordenadas x_i^*, así que de (25) resulta

$$f(\xi)=f(x_1^*\alpha_1+\ldots+x_n^*\alpha_n)=x_1^*f(\alpha_1)+\ldots+x_n^*f(\alpha_n)=x_1^*f(\alpha_1),$$

pues, por la definición del hiperplano, cada $f(\alpha_2), \ldots, f(\alpha_n)$ es cero. Como $\alpha_1, \alpha_2, \ldots, \alpha_n$ constituyen también una base ortogonal normal, podemos suponer que la constante $f(\alpha_1)$ es positiva. Por lo tanto, cualquier función lineal homogénea f puede ser expresada con relación a una conveniente base ortogonal normal en la forma $f(\xi) = dx_1^*$, donde $d > 0$.

Este resultado (en *pasiva*) puede ser traducido como en el Teorema 10, y entonces afirma que cualquier f lineal homogénea es equivalente mediante una transformación ortogonal $Y = XA$ a una función de la forma $g = dy_1$, con $d > 0$. En el grupo euclídeo también ésta es la forma canónica de las funciones f no homogéneas. Supongamos que $Y = XA + K$ es un movimiento rígido (con $A = \|a_{ij}\|$ ortogonal), el cual transforma a f en una de las formas $g = d'y_1$. Entonces, por sustitución, mediante (3) y (23), tenemos

$$b_1x_1 + \ldots + b_nx_n + c = d'y_1 = d'(x_1a_{11} + \ldots + x_na_{n1} + k_1).$$

Comparando los coeficientes nos resultará $b_i = d'a_{i1}$, así que $\Sigma_i b_i^2 = d'^2 \Sigma_i a_{i1}^2$. Pero la matriz A es ortogonal, luego $\Sigma_i a_{i1}^2 = 1$, según (11), y por tanto,

$$d'^2 = b_1^2 + b_2^2 + \ldots + b_n^2.$$

Si d' es positivo, esta relación determina unívocamente d' mediante los coeficientes b_i de la forma (23) dada. Por lo tanto, esta forma f es equivalente para el grupo euclídeo a sólo una de las formas $g = dy_1$ con $d > 0$, así que estas formas son canónicas y el coeficiente d que aparece en ellas es un invariante: no puede ser alterado por una transformación euclídea. La invariancia de d puede también probarse calculando directamente el efecto de cualquier transformación ortogonal, como en el siguiente Ejercicio 11. Estos hechos se resumen en el

TEOREMA 11. *En el grupo euclídeo, cualquier función lineal* $b_1x_1 + \ldots + b_nx_n + c$ *es equivalente a una, y sólo a una, de las formas canónicas* dy_1, *con* d *positivo; siendo, además,* $d = \sqrt{b_1^2 + \ldots + b_n^2}$ *invariante en este grupo.*

EJERCICIOS

1. Hallar formas canónicas de todos los polinomios *mónicos* cuadráticos $x^2 + bx + c$, en el grupo de las traslaciones.
2. Hallar formas canónicas de los polinomios $ax^2 + bx + c$ en el grupo afín.

3. En Ejerc. 2 mostrar que $d/a=b^2/a-4c$ es un invariante afín.
4. Demostrar que cualquier polinomio mónico de grado 3 es equivalente en las traslaciones a un polinomio en el que falta el término cuadrático.
5. Si $f(x)$ es cualquier polinomio de una variable, demostrar que el grado de f y el número de raíces reales son ambos invariantes en el grupo afín.
6. Para un polinomio con n variables, mostrar que los coeficientes de los términos de grado más elevado son invariantes en el grupo de las traslaciones.
* 7. Mostrar que cualquier polinomio cúbico real es equivalente bajo el grupo afín a una de las formas canónicas $a(x^3+x+c)$, $a(x^3-x+c)$, x^3+d, con $a>0$.
8. Hallar una forma canónica para las funciones lineales homogéneas, en el grupo de las semejanzas.
9. Tratar la misma cuestión: a) En el grupo diagonal de transformaciones $y_1=dx_1, \ldots, y_n=dx_n$; b) En el grupo monomial de todas las transformaciones $Y=XM$ con matriz monomial.
10. Demostrar que cualquier función $f(\xi)$ con las propiedades (25) es homogénea y lineal cuando se expresa mediante las coordenadas.
11. a) Demostrar que cualquier función lineal homogénea puede escribirse en forma de matriz, como $f=BX'$, donde B es la matriz de una fila $(b_1, \ldots, b_n)$.
    b) Utilizando las matrices, calcular el coeficiente c_i de la función transformada $g=c_1y_1+\ldots+c_ny_n$ equivalente a f en una transformación ortogonal $X=YA$.
    c) Demostrar, a partir de estas fórmulas, que $b_1^2+\ldots+b_n^2=c_1^2+\ldots+c_n^2$ (es decir, que d^2 es un invariante).
*12. Hallar las formas canónicas de la forma cuadrática x^2+bx+c en el grupo de traslaciones, cuando b y c son elementos en el campo J_2 de los enteros mód. 2.

6. Formas cuadráticas y matrices simétricas

Las cinco próximas secciones se dedicarán al estudio de las formas canónicas de las funciones cuadráticas, en distintos grupos de transformaciones. El más simple problema de este tipo aparece relacionado con el estudio de las cónicas planas con centro (elipse e hipérbolas referidas a ejes oblicuos). Tales cónicas tienen por ecuación $Ax^2+Bxy+Cy^2=1$, cuyo primer miembro es una «forma cuadrática». Estas formas cuadráticas (expresiones homogéneas de segundo grado en las variables) aparecen en muchas otras cuestiones: en las ecuaciones de las cuádricas del espacio, en la ecuación proyectiva de las cónicas en coordenadas homogéneas, en la fórmula para el cuadrado de la longitud de un vector $|X|^2=(x_1^2+x_2^2+ +\ldots+x_n^2)$, en la fórmula $(m/2)(u^2+v^2+w^2)$ de la energía cinética de un cuerpo moviéndose en el espacio con velocidad de componen-

tes u, v y w, y en geometría diferencial, al formular la longitud de un arco ds en coordenadas esféricas, $ds^2=dr^2+r^2d\phi^2+r^2 \operatorname{sen} \phi d\theta^2$, pudiéndose agregar muchos otros ejemplos.

Comencemos con unas nociones preliminares sobre matrices: una matriz A se llama *simétrica* si es igual a su transpuesta, $A'=A$; o, dicho de otro modo, $\|a_{ij}\|$ es simétrica si $a_{ij}=a_{ji}$ para i, j cualesquiera. Una matriz C se llama *hemisimétrica* si $C'=-C$. Para descomponer cualquier matriz B en una parte simétrica y otra hemisimétrica, escribiremos

$$B=(B+B')/2+(B-B')/2=S+K, \tag{26}$$

donde $S=(B+B')/2$, $K=(B-B')/2$. Por las reglas para calcular la transpuesta resulta $(B\pm B')'=B'\pm B''=B'\pm B$, así que S es simétrica y K hemisimétrica. No es posible ninguna otra descomposición del mismo tipo, $B=S_1+K_1$, con S_1 simétrica y K_1 hemisimétrica, pues tal descomposición daría $B'=S_1'+K_1'=S_1-K_1$, $B+B'=2S_1$, $B-B'=2K_1$ y $S_1=S$, $K_1=K$. Las fórmulas (26) se aplican siempre que $2=1+1\neq 0$, pero caen en defecto para las matrices sobre el campo de los enteros módulo 2, donde $1+1=0$. En conclusión, cualquier matriz puede expresarse de modo único como la suma de una matriz simétrica y otra hemisimétrica, supuesta que sea $1+1\neq 0$.

Para obtener una matriz a partir de una forma cuadrática, tal como $5x^2+6xy+2y^2$, escribamos ésta de modo que aparezcan los dos términos en xy e yx, como $5x^2+3xy+3xy+2y^2$. El resultado puede escribirse como producto de matrices

$$(x,\ y)\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=(x,\ y)\begin{pmatrix} 5x+3y \\ 3x+2y \end{pmatrix}=5x^2+6xy+2y^2.$$

El coeficiente 6 se ha repartido equitativamente entre xy e yx, así que la matriz 2×2 que resulta es simétrica.

Una interpretación matricial semejante se aplica a las formas en n variables. Una *forma cuadrática* (homogénea) es, por definición, cualquier polinomio $\sum_i \sum_j x_i b_{ij} x_j$ cuyos términos son todos de segundo grado. Esta forma puede escribirse así: XBX', como producto de matrices. Si la matriz B de los coeficientes es hemisimétrica, $b_{ij}=-b_{ji}$, y la forma es idénticamente nula. En general,

escribamos la matriz B como $B=S+K$, según (26); entonces la forma resulta

$$XBX'=X(S+K)X'=XSX'+XKX'=XSX' \qquad (K \text{ hemisimétrica})$$

Por lo tanto, si $1+1\neq 0$, cualquier forma cuadrática puede ser expresada unívocamente (ponemos A en vez de S) como sigue:

$$\text{(27)} \qquad \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j = XAX', \qquad A=\| a_{ij} \| \text{ simétrica.}$$

Si un vector ξ tiene coordenadas $X=(x_1, \ldots, x_n)$, cada forma cuadrática determina una función cuadrática $Q(\xi)=XAX'$ del vector ξ. Un cambio de base en el espacio dará las nuevas coordenadas X^* relacionadas con las antiguas por una ecuación $X=X^*P$, con P regular. Mediante las nuevas coordenadas de ξ, la función cuadrática resulta

$$Q(\xi)=XAX'=(X^*P)A(X^*P)'=X^*(PAP')X^{*\prime};$$

esto es, una nueva forma cuadrática de matriz PAP'. La nueva matriz es simétrica como A; $(PAP')'=P'A'P'=PAP'$.

TEOREMA 12. *Una transformación de coordenadas reemplaza una forma cuadrática de matriz* A *por una forma cuadrática de matriz* PAP', *siendo* P *regular.*

Paralelamente a esta interpretación «pasiva» está la correspondiente «activa»: toda transformación lineal no singular del espacio, $Y=XP^{-1}$, hace corresponder a la forma XAX' la forma $Y(PAP')Y'$.

Algunas veces se dice que dos matrices A y B son congruentes cuando (como en este caso) es $B=PAP'$ para alguna matriz P regular.

EJERCICIOS

1. Demostrar que $A'A$ y AA' son siempre simétricas.
2. Demostrar: si A es hemisimétrica, A^2 será simétrica.
3. Representar cada matriz de Ejerc. 1, § 2, Cap. VIII en la forma $S+K$.
4. Hallar la matriz simétrica asociada a cada una de las siguientes formas cuadráticas:
   a) $2x^2+3xy+6y^2$;
   b) $8xy+4y^2$;
   c) $x^2+2xy+4xz+3y^2+yz+7z^2$;
   d) $4xy$;
   e) $x^2+4xy+4y^2+2xz+z^2+4yz$.

5. a) Demostrar: si S es simétrica y A ortogonal, entonces $A^{-1}SA$ es simétrica.
   b) Si K es hemisimétrica y A es ortogonal, entonces $A^{-1}KA$ es hemisimétrica.
6. Estudiar la simetría de la matriz $AB-BA$ en los siguientes casos: a) A y B son simétricas ambas; b) A y B son ambas hemisimétricas; c) A es simétrica y B hemisimétrica.
7. Demostrar: si A y B son simétricas, entonces AB es simétrica si, y sólo si, $AB=BA$.
8. a) Demostrar: sobre el campo J_2 (enteros mód. 2) cualquier matriz hemisimétrica es simétrica.
   b) Presentar una matriz sobre J_2 que no sea una suma $S+K$ [crf. (26)].
9. Sean T y U transformaciones lineales sobre un espacio vectorial euclídeo con una base ortonormal dada. Demostrar que la matriz representando a U es la transpuesta de la matriz representando a T si, y sólo si, $(\xi T, \eta)=(\xi, \eta U)$ para cualquier par de vectores ξ y η. (Esta propiedad puede utilizarse para definir las transformaciones «simétricas» en un espacio euclídeo de infinitas dimensiones.)

7. Formas cuadráticas bajo el grupo lineal

El conocido proceso de «completar el cuadrado» puede utilizarse para la simplificación de una forma cuadrática mediante transformaciones lineales. Para dos variables, el proceso da

$$ax^2+2bxy+cy^2=a[x^2+2(b/a)xy+(b^2/a^2)y^2]+[c-(b^2/a)]y^2=$$
$$=a[x+(b/a)y]^2+[c-(b^2/a)]y^2.$$

Los términos entre corchetes sugieren el cambio de variables $x'=x+(b/a)y$, $y'=y$. La forma que resulta con esta transformación lineal es $ax'^2+[c-(b^2/a)]y'^2$; el término rectangular ha sido eliminado.

Este razonamiento exige que $a\neq 0$. Si $a=0$ pero $c\neq 0$, vale una transformación semejante. Finalmente, si $a=c=0$, la forma original es $2bxy$, y la correspondiente ecuación $2bxy=0$ representa una hipérbola equilátera. En este caso, la transformación $x=x'+y'$, $y=x'-y'$ reducirá la forma a

$$2b(x'+y')(x'-y')=2b(x'^2-y'^2);$$

este resultado contiene sólo los términos cuadrados. (*Sugerencia:* ¿Qué interpretación tiene la transformación empleada, respecto a los giros de los ejes de la hipérbola?)

Un método análogo de investigación preliminar puede aplicarse a las formas con más de dos variables.

LEMA. *Por una transformación lineal no singular, cualquier forma cuadrática* $\Sigma x_i a_{ij} x_j$ *no idénticamente nula, puede reducirse a una forma con primer coeficiente* $a_{11} \neq 0$, *excepto cuando* $1+1=0$.

Demostración. Por hipótesis, hay al menos un coeficiente $a_{ij} \neq 0$. Si éste es un término diagonal $a_{ii} \neq 0$, podemos convertirlo en el nuevo coeficiente $a_{11}' \neq 0$ por transposición de las dos variables x_1 y x_i (la cual es una transformación no singular, pues su matriz es una matriz de permutación). En otro caso, todos los términos diagonales serán cero, pero habrá dos índices $i \neq j$, con $a_{ij} \neq 0$. Permutando las variables podemos suponer $a_{12} \neq 0$. Por la simetría de la matriz es $a_{12}=a_{21}$. La forma cuadrática dada será, por tanto, $a_{12}x_1x_2+a_{21}x_2x_1+\ldots=2a_{12}x_1x_2$, más términos en los que aparecen otras variables. Lo mismo que en el caso de la hipérbola equilátera, la transformación

$$x_1=y_1-y_2, \quad x_2=y_1+y_2, \quad x_3=y_3, \quad \ldots, \quad x_n=y_n.$$

reduce lo precedente a la forma $2a_{12}(y_1^2-y_2^2)$, con el primer coeficiente $2a_{12} \neq 0$. Esta transformación es regular, pues por eliminación es fácil ver que tiene una transformación inversa

$$y_1=(x_1+x_2)/2, \quad y_2=(x_2-x_1)/2, \quad y_3=x_3, \quad \ldots, \quad y_n=x_n.$$

(Preguntamos: ¿en qué punto del razonamiento se utiliza la hipótesis de que $1+1 \neq 0$?)

Pasemos ahora a «completar el cuadrado» en una forma cuadrática. Según el lema, supondremos $a_{11} \neq 0$, así que la forma puede escribirse $a_{11}(\Sigma x_i b_{ij} x_j)$, donde $b_{ij}=a_{ij}/a_{11}$ y $b_{11}=1$. Por la simetría de la matriz, los términos en que aparece x_1 son

$$x_1^2+2\sum_{j=2}^{n} b_{1j}x_1x_j=\left(x_1+\sum_{j=2}^{n} b_{1j}x_j\right)^2-\left(\sum_{j=2}^{n} b_{1j}x_j\right)^2.$$

La formación de este «cuadrado perfecto» sugiere la transformación

$$y_1=x_1+\sum_{j=2}^{n} b_{1j}x_j, \quad y_2=x_2, \quad \ldots, \quad y_n=x_n;$$

con lo que y_1 aparece sólo en y_1^2. La forma original se ha convertido ahora en $a_{11}y_1^2+\Sigma y_j c_{jk} y_k$, donde los índices j y k varían desde 2 hasta n. Este sumatorio es, pues, una forma cuadrática en $n-1$ variables $y_2, \ldots, y_n$; a ésta se le puede aplicar el mismo proceso, y reiterarlo hasta que alguna de las formas cuadráticas residuales que van apareciendo, tenga nulos todos sus coeficientes. Por lo tanto, tenemos

TEOREMA 13. *Por una transformación lineal no singular de las variables, una forma cuadrática Q sobre un campo con $1+1\neq 0$, puede ser reducida a una forma diagonal cuadrática,*

(28) $$d_1y_1^2+d_2y_2^2+\ldots+d_ry_r^2, \quad \text{cada } d_i\neq 0.$$

El número r de los términos diagonales no nulos es un invariante llamado característica de la forma dada Q.

Sólo falta por demostrar la invariancia de r; esto es, hay que demostrar que la reducción de una Q dada a forma diagonal nos dará siempre el mismo número r de elementos no nulos. A este fin consideremos la forma como una función $Q(\eta)=d_1y_1^2+\ldots+d_ry_r^2$ del vector η de coordenadas $(y_1, \ldots, y_r, y_{r+1}, \ldots, y_n)$, con relación a una base $\epsilon_1, \ldots, \epsilon_n$, y busquemos aquellos vectores α que tienen la propiedad

(29) $$Q(\eta+\alpha)=Q(\eta), \quad \text{para todo } \eta \text{ en } V_n.$$

La expresión (28) de $Q(\eta)$ emplea sólo las r primeras coordenadas; por lo tanto, cualquier vector $\alpha=(0, \ldots, 0, a_{r+1}, \ldots, a_n)$ con sus primeras r coordenadas nulas cumple automáticamente la (29). Recíprocamente, si $\alpha=(a_1, \ldots, a_n)$ satisface a (29), será

$$Q(\eta+\alpha)-Q(\eta)=\Sigma d_i(y_i+a_i)^2-\Sigma d_iy_i^2=\Sigma d_ia_i^2+2\Sigma d_ia_iy_i=0.$$

Esto debe cumplirse para y_i cualesquiera. Luego con todas las $y_i=0$ quedará $\Sigma d_ia_i^2=0$; si sólo $y_j\neq 0$, dará $2d_ja_jy_j=0$, luego $a_j=0$ para cada $j=1, \ldots, r$. Por lo tanto, todos los vectores que cumplen (29) tienen $a_1=\ldots=a_r=0$.

Por lo tanto, los vectores α con la propiedad (29) forman un subespacio $(n-r)$ dimensional S, engendrado por la base $\epsilon_{r+1}, \ldots, \epsilon_n$. Pero la propiedad (29) define este subespacio de modo invariante, ya que es independiente del sistema particular de coordenadas.

Luego la dimensión $n-r$ de este subespacio es la misma en cualquier sistema de referencia. Esta dimensión (la *nulidad* de Q) determina a su vez la característica $r=n-(n-r)$.

La reducción de formas cuadráticas puede, por el Teorema 12, ser formulada como una reducción de las correspondientes matrices simétricas. Esto da el siguiente resultado:

COROLARIO 1. *Para cualquier matriz simétrica* A, *con elementos en un campo en que* $1+1\neq 0$, *existe una matriz regular* P *tal, que* PAP′ *es diagonal. El número de elementos diagonales no nulos en* PAP′ *es el mismo para todas las matrices diagonales obtenidas así a partir de la* A.

Si la característica r de la forma cuadrática es igual a n, número de las variables, la matriz diagonal $D=PAP'$ tendrá todos los elementos diagonales no nulos, luego será regular. Asimismo $A=P^{-1}D(P')^{-1}$, por ser un producto de matrices regulares, será regular. Con esto demostramos

COROLARIO 2. *Una forma cuadrática* XAX′ *con* n *variables* X, *tiene característica* n *cuando la matriz* A *es regular, y sólo en este caso.*

Por este motivo, una forma cuadrática de característica n es llamada regular o no singular. En todo caso, solamente el número r de términos de la forma (28) está determinado, pero no lo están sus coeficientes. Distintos modos de reducción de la forma pueden dar diferentes conjuntos de coeficientes. Por este motivo, la forma (28) no es única (o canónica) en un campo general F de coeficientes. Veamos ahora las especiales circunstancias que se presentan en un campo F particular, esto es, en el campo real.

EJERCICIOS

1. Sobre el campo de los números racionales, reducir cada forma cuadrática del § 6, Ejerc. 4, a forma diagonal.
2. Reducir $2x^2+xy+3y^2$ a forma diagonal sobre el campo de enteros mód. 5.
3. Sobre el campo de enteros mód. 5 demostrar que cualquier forma cuadrátrica puede reducirse por transformaciones lineales a la forma $\Sigma d_i y_i^2$, con cada coeficiente $d_i=0$, 1 ó 2.
4. Sobre el campo de los números racionales, mostrar que la forma cuadrática $x_1^2+x_2^2$ puede transformarse en las dos formas diagonales distintas $9y_1^2+4y_2^2$ y $2z_1^2+8z_2^2$.

5. Hallar una P tal, que PAP sea diagonal, siendo

a) $A=\begin{pmatrix}3 & 1\\ 1 & 0\end{pmatrix}$. b) $A=\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$. c) $A=\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 2\\ 0 & 2 & 0\end{pmatrix}$.

6. Hallar todas las transformaciones lineales que transforman la forma real cuadrática $x_1^2+\ldots+x_n^2$ en $y_1^2+\ldots+y_n^2$.

8. Formas cuadráticas reales bajo el grupo lineal

En la geometría analítica se describen las secciones cónicas mediante formas cuadráticas cuyos coeficientes pertenecen al campo de los números reales. En este campo, cada coeficiente d_i de una forma diagonal $\Sigma d_i y_i^2$ puede simplificarse más, mediante la sustitución $y_i'=d_i^{1/2}y_i$, con la que el término $d_i y_i^2$ se reduce exactamente a $y_i'^2$. Esta sustitución puede hacerse sólo cuando $d_i>0$, condición precisa para el cálculo de la raíz cuadrada de d_i. Si $d_i<0$, no tiene raíz cuadrada real. En este caso sustituiremos $y_i'=(-d_i)^{1/2}y_i$, con lo que $d_i y_i^2$ se convierte en $-y_i'^2$. Haciendo la transformación análoga con todas las variables podremos reducir la forma cuadrática al tipo $\Sigma \pm y_i^2$. En esta suma podemos permutar las variables de modo que los cuadrados positivos sean los primeros. Esto demuestra

TEOREMA 14. *Cualquier forma cuadrática* Q *sobre el campo de los números reales, puede ser reducida por una transformación lineal no singular de las variables, a la forma*

$$(30)\qquad Q(\xi)=z_1^2+\ldots+z_p^2-z_{p+1}^2-\ldots-z_r^2.$$

TEOREMA 15. *El número* p *de cuadrados positivos que aparecen en la forma reducida (30), es un invariante de la forma dada; es decir que depende sólo de la forma, y no del método seguido para reducirla (Ley de inercia, de Sylvester).*

Demostración. Imaginemos que existiese otra forma reducida

$$(31)\qquad Q(\xi)=y_1^2+\ldots+y_q^2-y_{q+1}^2-\ldots-y_r^2,$$

con q términos positivos. Como ambas son obtenidas de Q por transformaciones lineales regulares, (30) y (31), podrían transformarse una en otra por una transformación regular. Podemos considerar las ecuaciones de esta última transformación como un cambio de coordenadas (por pasiva); entonces (30) y (31) representan

la misma función cuadrática $Q(\xi)$ de un vector con coordenadas z_i relativas a una base, y coordenadas y_i relativas a otra base.

Supongamos que $q<p$. Entonces $Q(\xi)\geqslant 0$ siempre que $z_{p+1}=$ $=\ldots=z_r=0$ en (30). Las ξ que satisfacen a estas $r-p$ ecuaciones forman un subespacio S_1 de $n-(r-p)$ dimensiones (en este subespacio, las $n-(r-p)$ coordenadas son $z_1, \ldots, z_p, z_{r+1}, \ldots, z_n$). De un modo semejante, (31) da $Q(\xi)<0$ para cada $\xi\neq 0$ con coordenadas $y_1=\ldots=y_q=y_{r+1}=\ldots=y_n=0$. Estas condiciones determinan un subespacio S_2 de $r-q$ dimensiones. La suma de las dimensiones de los espacios S_1 y S_2 es $n-(r-p)+(r-q)=n+(p-q)>n$. Por lo tanto, S_1 y S_2 tienen algún vector ξ no nulo en común, pues según el Teorema 9 de Cap. VII, la dimensión de la intersección $S_1 \frown S_2$ es positiva. Para este vector común ξ, será $Q(\xi)\geqslant 0$, por (30), y $Q(\xi)<0$, por (31), lo que es manifiestamente contradictorio. La suposición $q>p$ lleva a otra contradicción semejante, así que $q=p$, como queríamos demostrar.

Este resultado nos enseña que cualquier forma real cuadrática puede ser reducida mediante transformaciones lineales a una, y sólo una, forma del tipo (30). Las expresiones $\Sigma \pm z_i^2$ de este tipo son, pues, las formas cuadráticas *canónicas* en el grupo lineal. En cuanto a esta misma forma canónica está determinada unívocamente por la llamada *signatura* $\{+, \ldots, +, -, \ldots, -\}$ que es una sucesión de p signos positivos y $r-p$ signos negativos, siendo r la característica de la forma.

El conjunto de signos está determinado por r y por la diferencia $s=p-(r-p)=2p-r$ entre el número de signos positivos y el de negativos. Algunas veces se llama signatura de la forma al entero s. En todo caso r y s forman un sistema completo de invariantes numéricos, puesto que dos formas son equivalentes si, y sólo si, se reducen ambas a la misma forma canónica (30).

TEOREMA 16. *Dos formas cuadráticas reales son equivalentes para el grupo lineal si, y sólo si, ambas tienen la misma característica y la misma signatura.*

Una forma cuadrática Q en n variables se llama *definida positiva* si su forma canónica $z_1^2+\ldots+z_n^2$ consta de n cuadrados positivos. Como una suma de cuadrados de números reales sólo es cero si lo es cada uno de los términos, una Q definida positiva toma un valor $Q(\xi)>0$ siempre que $\xi\neq 0$. No es éste el caso cuando en la

forma canónica (3) faltan algunos términos (30) o aparecen algunos términos negativos. Por otra parte, la propiedad $Q(\xi) > 0$ es, naturalmente, independiente de las coordenadas empleadas para determinar ξ; lo mismo se aplica a la forma original $Q = XAX'$.

TEOREMA 17. *Una forma real cuadrática* $\Sigma x_i a_{ij} x_j$ *es definida positiva si, y sólo si,* $\Sigma x_i a_{ij} x_j > 0$, *excepto para* $x_1 = \ldots = x_n = 0$.

Una matriz simétrica A se llama *definida positiva* cuando su forma correspondiente XAX' es definida positiva. Por definición, esto sólo ocurre cuando la forma es equivalente a la canónica $z_1^2 + \ldots + z_n^2$ con I como matriz. Por el Teorema 12, esto significa que $A = PIP'$, y tendremos el siguiente resultado:

TEOREMA 18. *Una matriz real simétrica* A *es definida positiva si, y sólo si, existe una matriz regular y real* P *tal, que* $A = PP'$.

Una forma cuadrática XAX' determina en un espacio n-dimensional una figura, lugar de los puntos X tales, que $XAX' = 1$. La forma canónica (30) significa que una conveniente transformación lineal regular puede reducir este lugar al de ecuación

$$z_1^2 + \ldots + z_p^2 - z_{p+1}^2 - \ldots - z_r^2 = 1.$$

Por ejemplo, en el plano, las ecuaciones reducidas de características 2 son

$$x^2 + y^2 = 1, \qquad x^2 - y^2 = 1, \qquad -x^2 - y^2 = 1,$$

que representan respectivamente un círculo, una hipérbola equilátera y un lugar vacío (esto es, sin puntos reales). La única forma de característica 0 da $0 = 1$. Las únicas ecuaciones de características 1 son $x^2 = 1$ (que representa las dos líneas $x = \pm 1$) y $-x^2 = 1$ (lugar vacío). En el próximo capítulo demostraremos (Teorema 9) que cualquier transformación lineal no singular del plano puede ser representada como producto de corrimientos, compresiones y reflexiones. Por lo tanto, cualquier cónica con centro $ax^2 + bxy + cy^2 = 1$ puede ser reducida a una de las formas enumeradas por una sucesión de corrimientos, compresiones y reflexiones. Geométricamente, este resultado es razonable; una elipse puede comprimirse según la dirección del eje mayor hasta dar un círculo; pero, naturalmente, no habrá transformación lineal que reduzca al círculo $x^2 + y^2 = 1$ la

hipérbola equilátera $x^2 - y^2 = 1$. Tal es el significado geométrico de la invariancia de la signatura.

La signatura es también utilizable al estudiar el máximo o mínimo de las funciones de dos variables. Sea $z = f(x, y)$ una función cuyas derivadas primeras f_x' y f_y' se anulan para $x = x_0$, $y = y_0$. Por lo tanto, en el desarrollo de Taylor de z en potencias de $h = (x - x_0)$ y $k = (y - y_0)$, faltarán los términos de primer grado. Este desarrollo (supuesto convergente) es

$$f(x_0 + h,\ y_0 + k) = f(x_0,\ y_0) + (1/2)[ah^2 + 2bhk + ck^2] + \ldots;$$

siendo los coeficientes

$$a = f_{xx}''(x_0,\ y_0), \quad b = f_{xy}''(x_0,\ y_0), \quad c = f_{yy}''(x_0,\ y_0).$$

Para pequeños valores de h y de k, los términos importantes son los del corchete; en él se encuentra una forma cuadrática en h y k con coeficientes reales. Si esta forma tiene característica 2, puede expresarse mediante unas nuevas variables convenientes, como $\pm h'^2 \pm k'^2$. Si ambos signos son más, los valores de $f(x_0 + h, y_0 + k)$ excederán siempre al de $f(x_0, y_0)$ y z presentará un *mínimo* relativo. Si los dos signos son menos, z tendrá un *máximo*. Si un signo es más y otro menos, la forma cuadrática puede tomar valores positivos o negativos, así que (x_0, y_0) no es ni un máximo ni un mínimo, sino un llamado punto de silla. Estos distintos puntos notables de f son distinguidos por la signatura de la forma cuadrática. Los resultados son análogos en el estudio de los puntos notables en las funciones de tres o más variables.

EJERCICIOS

1. Demostrar: una forma cuadrática homogénea con coeficientes complejos es siempre equivalente en el grupo lineal (utilizando coeficientes complejos) a una suma de cuadrados $z_1^2 + \ldots + z_r^2$.
2. Demostrar que dos formas cuadráticas en n variables con coeficientes complejos, son equivalentes en el grupo lineal si, y sólo si, tienen ambas la misma característica.
3. Reducir las siguientes formas cuadráticas reales a la forma canónica de Teorema 14. Hallar la característica, nulidad y signatura de cada forma.

   a) $9x_1^2 + 12x_1x_2 + 79x_2^2$;
   b) $2x_1^2 - 12x_1x_2 + 18x_2^2$;
   c) $-2x_1^2 - 4x_1x_2 + 22x_2^2 + 12x_2x_3 + 6x_3x_1 - x_3^2$.

4. Describir los lugares geométricos que corresponden a los varios casos posibles de forma canónica de las formas cuadráticas reales en tres dimensiones.
5. a) Enumerar todos los tipos de formas cuadráticas no singulares en cuatro variables.
   b) Describir geométricamente dos por lo menos de los lugares correspondientes en V_4.
6. Demostrar que $ax^2+bxy+cy^2$ es definida positiva si, y sólo si, $4ac-b^2>0$.
7. Una forma cuadrática se llama *semidefinida positiva* si su característica es igual a su signatura. Enunciar y demostrar para estas formas un teorema análogo al 17.
8. Hacer lo mismo para el Teor. 18.

9. Formas cuadráticas bajo el grupo ortogonal

¿Cómo puede simplificarse una forma cuadrática real mediante transformaciones ortogonales? Una transformación ortogonal $Y=XP$ cambia XAX' en $Y(P^{-1}AP^{-1'})Y'$; por ser P ortogonal, la nueva matriz puede escribirse (*) $P^{-1}AP^{-1'}=P^{-1}AP$.

En el plano, una transformación ortogonal (rotación o reflexión) de una elipse no puede dar nunca un círculo. A lo más se podrán llevar los ejes de la elipse a una posición típica. El eje focal viene caracterizado como el mayor de los diámetros. A esta propiedad de máximo se le puede dar otra forma; para hacerlo, consideremos una función real cuadrática $Q(\xi)=ax^2+cy^2$ con $a\leqslant c$ y sin término en xy. Entonces $Q(\xi)\leqslant cx^2+cy^2=c(x^2+y^2)$: esto significa que el valor máximo de Q para todos los puntos del círculo unidad $x^2+y^2=1$ es c; este máximo lo alcanza Q en el punto $y=1$, $x=0$ Inversamente, este último hecho asegura la falta en Q del término en xy.

LEMA. *Si una función cuadrática real* $Q=ax^2+bxy+cy^2$ *calculada sobre los puntos del círculo* $x^2+y^2=1$ *presenta un máximo para* $x=0$, $y=1$, *debe ser* $b=0$.

Demostración. Q es una función (bivalente) de la variable x, donde y está dada implícitamente por $x^2+y^2=1$. Derivando ésta, $2x+2y(dy/dx)=0$, así que $y'=dy/dx$ es $y'=-x/y$. La derivada de Q será, pues,

$$Q'=(ax^2+2bxy+cy^2)'=2ax+2by+2bxy'+2cyy'.$$

(*) Dos matrices simétricas A y $P^{-1}AP$, con P ortogonal, son llamadas a veces *ortogonalmente congruentes*. Ambas son equivalentes en el grupo de todas las transformaciones $A\to P^{-1}AP$ del conjunto de matrices A.

Sustituyendo el valor de y' y poniendo $y=1$, $x=0$, resulta $Q'=2b$, pero en el máximo, $y=1$, $x=0$, la derivada debe anularse, luego $2b=0$, c. q. d.

Admitamos ahora que cualquier forma real cuadrática $Q(\xi)= = \sum_i \sum_j x_i a_{ij} x_j$ presente un máximo λ_1 para el conjunto de los puntos que pertenecen a la *hiperesfera unidad* $\Sigma x_i^2=1$. Esto es, que entre todos los vectores ξ de longitud unidad, hay uno ξ_0 en el cual $Q(\xi)$ toma su valor máximo λ_1 (*). Como la longitud de ξ_0 es 1, podemos tomar $\alpha_1=\xi_0$ como primer vector de una nueva base normal ortogonal $\alpha_1, \ldots, \alpha_n$ (Lema 2, § 9, Cap. VII). En función de las nuevas coordenadas $y_1, \ldots, y_n$ de ξ relativas a esta base, la forma cuadrática se expresa ahora como $Q(\xi)=\Sigma y_i b_{ij} y_j$ con una nueva matriz de coeficientes b_{ij}. El valor λ_1 máximo de Q está dado por el vector α_1 con coordenadas $(1, 0, \ldots, 0)$; luego, por sustitución, vemos que este valor máximo es b_{11}. Ahora bien, el mismo máximo obtendremos aunque nos limitemos a considerar sólo aquellos vectores cuyas coordenadas son nulas excepto y_1 y otra de ellas, y_i por ejemplo. Por lo tanto, $y_1=1$, $y_i=0$ será el máximo de la forma $b_{11}y_1^2+2b_{1i}y_1y_i+b_{ii}y_i^2$, con la condición $y_1^2+y_i^2=1$. El lema (con x reemplazado por y_i) demuestra que en tal caso el coeficiente b_{1i} es cero. Aplicando este razonamiento a los casos $i=2, \ldots, n$ resulta que en la expresión de Q mediante las coordenadas y_i deben faltar todos los términos rectangulares en que intervenga y_1 y por lo tanto

$$(32) \qquad Q(\xi)=\lambda_1 y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n} y_i b_{ij} y_j, \qquad B=\| b_{ij} \|=B'.$$

El primer coeficiente λ_1 *no* es un vector, sino un escalar (el máximo de $Q(\xi)$ para $|\xi|=1$).

La diferencia $Q^*(\xi)=Q(\xi)-\lambda_1 y_1^2$ igual al sumatorio que aparece en (32), es también una forma cuadrática, con las $n-1$ variables $y_2, \ldots, y_n$. Estas variables son coordenadas del espacio S_{n-1} engendrado por los $(n-1)$ vectores básicos $\alpha_2, \ldots, \alpha_n$. En este espacio (que es el complemento ortogonal del vector ξ_0) puede ser aplicado el mismo proceso para escoger una nueva base ortogonal,

(*) En el cálculo elemental se da por supuesto que existe el máximo λ_1. Una prueba sencilla de que en efecto es así, puede ser dada en estas líneas: Para $\Sigma x_i^2=1$, Q es acotado, luego tiene un extremo superior λ_1 (por la propiedad básica de los números reales, Cap. III). Como Q es continua y la hiperesfera es un conjunto cerrado, el extremo superior debe ser accesible con algún vector ξ_0.

produciendo el máximo de $Q^*(\xi)$ para $|\xi|=1$; esto destaca otro término diagonal de la forma. Finalmente, se encuentra una base de *ejes principales*, para la cual es

$$(33) \qquad Q(\xi)=\lambda_1 z_1{}^2+\lambda_2 z_2{}^2+\ldots+\lambda_n z_n{}^2, \quad \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n.$$

Sean $z_1, \ldots, z_n$ las coordenadas de ξ relativas a la base $\alpha_1, \beta_2, \gamma_3, \ldots$, la cual ha sido construída paso a paso por sucesivas investigaciones de máximo. El primer vector α_1 da para $Q(\xi)$ un máximo λ_1 condicionado por $|\xi|=1$. El segundo vector básico β_2 es un vector elegido en el espacio ortogonal a α_1; esto es, $\eta=\beta_2$ dará para $Q(\eta)$ un máximo λ_2 entre los vectores η para los cuales $|\eta|=1$ y $(\eta, \alpha_1)=0$. El tercer vector de la base γ_3 da un máximo para $Q(\zeta)$ entre todos los vectores $|\zeta|=1$ ortogonales a α_1 y β_2; y así sucesivamente. Estos sucesivos problemas de máximo pueden visualizarse (en forma inversa) sobre un elipsoide con tres ejes distintos $a>b>c>0$. El menor de los ejes principales c es el diámetro mínimo. El siguiente eje principal b es el diámetro mínimo de todos los que son perpendiculares al antedicho c, etc.

Los coeficientes λ_i de (33) pueden, pues, caracterizarse como soluciones de ciertos problemas de máximo que dependen sólo de Q y no del sistema de coordenadas. El único caso en que es posible una ambigüedad en el proceso de reducción, es cuando el primer máximo, o alguno de los sucesivos, es alcanzado por dos ó más vectores distintos ξ_0 y η_0 de longitud unidad. También en este caso se puede demostrar que λ_i es única.

Según el Corolario del Teorema 10, el proceso de reducir Q por elección de una nueva base ortogonal normal de coordenadas, es equivalente a la reducción de Q por transformación ortogonal de los variables, y por lo tanto,

TEOREMA 19. *Por una transformación ortogonal de variables, cualquier forma cuadrática real puede ser reducida a la forma diagonal (33).*

En el Cap. X, § 9, demostraremos que esta forma es canónica, es decir, que Q es equivalente por el grupo ortogonal a una, y sólo a una, de tales formas. Además, el cálculo explícito de la transformación ortogonal que reduce Q a esta forma canónica, se lleva a cabo del modo más eficiente con el empleo de las raíces características, discutidas en dicho capítulo.

En el plano, las formas canónicas son simplemente $\lambda_1 x^2 + \lambda_2 y^2$; con ellas se relaciona la usual ecuación típica de una elipse ($\lambda_1 \geqslant \lambda_2 \geqslant 0$) y la de la hipérbola ($\lambda_1 > 0$, $\lambda_2 < 0$); los coeficientes determinan la longitud de los ejes. En el espacio tridimensional, una observación análoga se aplica a los tres coeficientes λ_1, λ_2, λ_3. Si son los tres positivos, el lugar $Q = 1$ es un elipsoide. Si uno es negativo, un hiperboloide de una hoja; si dos son negativos, un hiperboloide de dos hojas. Si los tres son negativos, un lugar vacío. Se notará en todo caso la significación de la signatura y de la característica.

EJERCICIOS

1. Demostrar que para cualquier matriz real simétrica A, existe una matriz ortogonal P tal, que $P^{-1}AP$ es diagonal.
2. Demostrar que cualquier matriz hemisimétrica A tiene la forma $A = P^{-1}BP$, con P ortogonal y B^2 diagonal.
3. Reducir las siguientes formas cuadráticas a forma diagonal por transformación ortogonal, siguiendo el método dado:

   a) $5x^2 - 6xy + 5y^2$; b) $2x^2 + 4\sqrt{3}\,xy - 2y^2$.
4. A la forma cuadrática $9x_1^2 - 9x_2^2 + 18x_3^2$ se aplica la transformación ortogonal:

   $3x_1 = 2y_1 - y_2 + 2y_3$, $3x_2 = -y_1 + 2y_2 + 2y_3$, $3x_3 = 2y_1 + 2y_2 - y_3$.

   Para la forma Q en y_1, y_2, y_3 que así resulta, demostrar directamente que el vector (2/3, 2/3, —1/3) dará el máximo valor 18 para Q cuando

   $$y_1^2 + y_2^2 + y_3^2 = 1.$$

   Comprobarlo por el cálculo.
5. * Si el vector (1, 0, ..., 0) da el valor máximo para $\sum_{ij} x_i a_{ij} x_j$, condicionado por $\sum_i x_i^2 = 1$, demostrar por el método de los multiplicadores de Lagrange que $a_{1j} = a_{j1} = 0$ para $j \neq 1$.
6. a) Demostrar que las coordenadas x_0, y_0 del vector que reduce a un máximo la forma cuadrática $ax^2 + 2bxy + cy^2$, con la condición $x^2 + y^2 = 1$, satisface a $y_0/x_0 = [(c-a) \pm \sqrt{(a-c)^2 + 4b^2}]/2b$, si $b \neq 0$.

   b) Demostrar que las dos elecciones de signo en esta fórmula determinan vectores ortogonales.

10. Cuádricas bajo los grupos afín y euclídeo

Consideremos ahora una función cuadrática *no homogénea* arbitraria de un vector de coordenadas $x_1, \ldots, x_n$,

$$(34) \qquad f(\xi) = \sum_i \sum_j x_i a_{ij} x_j + \sum_k b_k x_k + c, \qquad (i,\ j,\ k = 1,\ \ldots,\ n).$$

La misma puede escribirse $f(\xi)=XAX'+BX'+c$, donde $A=\| a_{ij} \|$ es una matriz simétrica y $B=(b_1, \ldots, b_n)$ una matriz de una fila.

En el caso más sencillo de una función de una variable $f=ax^2+$ $+bx+c$ se observa que una traslación $x=y+k$ deja inalterado el coeficiente a del cuadrado, pues

$$(35) \qquad f=a(y+k)^2+b(y+k)+c=ay^2+(2ak+b)y+ak^2+bk+c.$$

Un cálculo análogo vale para n variables; una traslación $X \rightarrow Y=X-K$ (K matriz de una fila) da

$$f(\xi)=(Y+K)A(Y+K)'+B(Y+K)'+c=$$
$$=YAY'+KAY'+YAK'+KAK'+BY'+BK'+c.$$

El producto YAK' (matriz fila $\times$ matriz $\times$ matriz columna) es un escalar, luego también lo es su transpuesta $KA'Y'=KAY'$; por tanto,

$$(36) \qquad f(\xi)=YAY'+(2KA+B)Y'+KAK'+BK'+c$$

fórmula exactamente análoga a la (35). Esto demuestra que

LEMA. *En cualquier traslación permanece invariable la matriz* A *de la parte homogénea de segundo grado de una función cuadrática* f(ξ).

Por otra parte, una transformación lineal homogénea $X=YP$ cambiará $f(\xi)$ en $Y(PAP')Y'+(BP')Y'+c$; en esta función cuadrática, la nueva matriz de los términos cuadráticos es PAP', exactamente como en el caso de la transformación de una forma homogénea.

¡Se trata ahora de reducir la función real $f(\xi)$ por un movimiento rígido, de ecuaciones $X=YP+K$ (P ortogonal)! Según la observación anterior, la transformación ortogonal P se empleará sólo para simplificar la matriz A de los términos cuadráticos, exactamente como para una forma cuadrática homogénea. Como en § 9, encontramos (con nuevos coeficientes b_i')

$$f(\xi)=\lambda_1 z_1{}^2+\ldots+\lambda_n z_n{}^2+b_1'z_1+\ldots+b_n'z_n+c.$$

Las b_j' asociadas con las λ_j *no nulas* pueden eliminarse ahora por el simple método de «completar los cuadrados», como en (22)

de § 5, mediante la traslación $y_i = z_i + b_i'/2\lambda_i$. Ahora, permutando las variables para que las λ_i no nulas figuren las primeras, obtenemos

$$f(\xi) = \lambda_1 y_1^2 + \dots + \lambda_r y_r^2 + b_{r+1}' z_{r+1} + \dots + b_n' z_n + c'.$$

Si la parte lineal de esta función no es precisamente la constante c', puede transformarse en $d y_{r+1}$ mediante una conveniente traslación y transformación ortogonal, como en § 5, Teor. 11. Esta transformación no afecta en absoluto a las r primeras variables.

El resultado es una de las formas

$$(37) \qquad f(\xi) = \lambda_1 y_1^2 + \dots + \lambda_r y_r^2 + d y_{r+1},$$

$$(38) \qquad f(\xi) = \lambda_1 y_1^2 + \dots + \lambda_r y_r^2 + c',$$

donde $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_r$, ninguna $\lambda_i = 0$, $d > 0$.

TEOREMA 20. *En el grupo euclídeo de los movimientos rígidos, toda forma cuadrática (34) es equivalente a una de las dos formas (37) o (38).*

Estas formas reducidas que acabamos de obtener son canónicas en el grupo de los movimientos rígidos. Para demostrarlo, admitamos el resultado, que será probado en Cap. X, de que la forma diagonal del Teorema 19 para una forma homogénea, es única. Sigue de aquí que, para las formas (37) y (38), la característica r y los coeficientes $\lambda_1, \dots, \lambda_r$ están determinados unívocamente, pues son, simplemente, la característica y la diagonal de los coeficientes que corresponden a la matriz simétrica A, y estas cantidades son invariantes en las transformaciones ortogonales $A \to PAP'$ a que se somete A. Las cantidades d y c' son también invariantes. Se trata, precisamente, de los invariantes de la función lineal $g(z) = b_{r+1}' z_{r+1} + \dots + b_n' z_n + c'$, y esta función lineal está unívocamente determinada como el valor de f sobre el subespacio para el que es $y_1 = \dots = y_r = 0$. Este subespacio puede describirse, en términos invariantes, como el conjunto de todos los vectores α tales, que $f(\xi + \alpha) = f(\xi) + f(\alpha)$ para cualquier vector ξ.

Para las transformaciones afines $X = YP + K$, con P regular, se aplica un razonamiento análogo. Reduciendo la parte cuadrática a forma diagonal, los coeficientes pueden ahora ser tomados iguales a ± 1, como en § 8. La parte lineal se tratará después como en § 5.

TEOREMA 21. *Por una transformación afín (o por un cambio afín de coordenadas) cualquier función cuadrática real en* n *variables puede reducirse a una de las formas*

$$(39)\qquad y_1^2+\ldots+y_p^2-y_{p+1}^2-\ldots-y_r^2+y_{r+1},\qquad (r<n),$$

$$(40)\qquad y_1^2+\ldots+y_p^2-y_{p+1}^2-\ldots-y_r^2+c,\qquad (r\leqslant n).$$

Como los términos cuadráticos no son alterados por la traslación, la característica r y el número p de términos positivos deben ser invariantes, por la ley de inercia (Teorema 15).

EJERCICIOS

1. Clasificar en el grupo euclídeo las formas
   a) $4xz+4y^2+8y+8$;
   b) $9x^2-4xy+6y^2+3z^2+2\sqrt{5}x+4\sqrt{5}y+12z+16$.
2. Clasificar en el grupo afín las formas
   a) $x^2+4y^2+9z^2+4xy+6xz+12yz+8x+16y+24z+15$;
   b) $x^2-6xy+10y^2+2xz-20z^2-10yz-40z-17$;
   c) $x^2+4z^2+4xz+4x+4z-6y+6$;
   d) $-2x^2-3y^2-7z^2+2xy-8yz-6xz-4x-6y-14z-6$.
3. En la función cuadrática $XAX'+BX'+c$, con la matriz A regular, demostrar que los términos lineales pueden suprimirse por una traslación.
4. Calcular el efecto de una transformación arbitraria afín, sobre una función cuadrática $XAX'+BX'+c$.
5. Generalizar la clasificación afín de la función cuadrática dada en Teorema 21, a las funciones con coeficientes en cualquier campo en el que $1+1\neq 0$.

* 11. Matriz unitaria, matriz hermítica

Pasando ahora al campo de los números complejos, reemplazaremos las precedentes transformaciones ortogonales de las formas cuadráticas reales por las transformaciones «unitarias» de ciertas formas «hermíticas». Un número complejo $c=a+ib$ se define como un par ordenado de números reales (a, b) o como un vector de componentes (a, b) en el espacio de *dos* dimensiones $V_2(R^*)$. La norma o valor absoluto $|c|$ del complejo es precisamente la longitud de este vector real

$$(41)\qquad |c|^2=|a+ib|^2=a^2+b^2=(a+ib)(a-ib)=cc^*,$$

donde c^* denota el complejo conjugado $a-ib$. Con el mismo fundamento, un vector γ con n componentes complejos $(c_1, \ldots, c_n)$, donde $c_j=a_j+ib_j$, puede considerarse como un vector con $2n$ componentes $(a_1, \ldots, a_n, b_1, \ldots, b_n)$ en un espacio real de doble número de dimensiones. La longitud de este vector real está dada por

$$(42) \qquad |(c_1, \ldots, c_n)|^2=(a_1^2+b_1^2)+\ldots+(a_n^2+b_n^2)= \\ = \sum_{j=1}^{n}(a_j+ib_j)(a_j-ib_j)=c_1c_1^*+\ldots+c_nc_n^*$$

Como cada producto $c_jc_j^*=a_j^2+b_j^2 \geqslant 0$, esta expresión tiene la propiedad esencial de ser *definida positiva:* la suma real $\sum_{j=1}^{n} c_jc_j^*$ es positiva, excepto cuando todas las c_j son nulas. En este aspecto, (42) se asemeja a la usual expresión pitagórica de la longitud de un vector *real*. Así pues, adoptaremos (42) como definición de la longitud del vector complejo $K=(c_1, \ldots, c_n)$. La fórmula $\Sigma c_jc_j^*$ puede escribirse en notación matricial como producto $KK^{*\prime}$ donde K^* es el vector cuyas componentes son las conjugadas de las de K.

DEFINICIÓN. *En un espacio vectorial* $V_n(C)$ *sobre el campo* C *de los números complejos, sean* ξ *y* η *dos vectores con coordenadas* $X=(x_1, \ldots, x_n)$ *e* $Y=(y_1, \ldots, y_n)$ *respectivamente; introduzcamos un producto interno, mediante la fórmula*

$$(43) \qquad (\xi, \eta)=x_1y_1^*+\ldots+x_ny_n^*=XY^{*\prime}$$

La longitud de ξ *es entonces* $|\xi|=(\xi, \xi)^{1/2}$.

Como en el caso del ordinario producto interno pueden demostrarse las propiedades básicas

Lineal: $\qquad (c\xi+d\eta, \zeta)=c(\xi, \zeta)+d(\eta, \zeta)$,

Semisimétrica: $\qquad (\xi, \eta)=(\eta, \xi)^*$,

Positiva: $\qquad$ Si $\xi\neq 0$, (ξ, ξ) es real y positivo.

La semisimetría implica claramente una semilinealidad en el segundo factor: $(\xi, c\eta+d\zeta)=(c\eta+d\zeta, \xi)^*=c^*(\eta, \xi)^*+d^*(\zeta, \xi)^*= =c^*(\xi, \eta)+d^*(\xi, \zeta)$, así que

$$(44) \qquad (\xi, c\eta+d\zeta)=c^*(\xi, \eta)+d^*(\xi, \zeta).$$

Si se desea, pueden adoptarse las propiedades lineal, semisimétrica y positiva como la definición axiomática del producto interno (ξ, η) en un espacio vectorial abstracto sobre el campo de los números complejos; el espacio se llama entonces *espacio unitario* (comparar con el espacio vectorial euclídeo abstracto de Cap. VII, §8).

Dos vectores ξ y η serán *ortogonales* ($\xi \perp \eta$) cuando sea $(\xi, \eta)=0$. Por la semisimetría, $\xi \perp \eta$ implica $\eta \perp \xi$. Un conjunto de n vectores $\alpha_1, \ldots, \alpha_n$ en el espacio (n-dimensional) será una base *unitaria normal* del espacio, si cada vector tiene longitud unidad y si cada dos son ortogonales:

$$(45) \qquad |\alpha_1| = \ldots = |\alpha_n| = 1, \qquad (\alpha_i, \alpha_j) = 0 \qquad (i \neq j).$$

Tal conjunto de vectores es evidentemente una base, en el sentido ordinario de la palabra. Los vectores de la base original $\epsilon_1 = (1, 0, \ldots, 0), \ldots, \epsilon_n = (0, \ldots, 0, 1)$ constituyen una de estas bases; y como en el §9 del Cap. VII, se pueden construir bases distintas y demostrar

TEOREMA 22. *Cualquier conjunto de* $m < n$ *vectores mutuamente ortogonales y de longitud uno en un espacio unitario, forman parte de una base unitaria normal del espacio.*

En particular, cualquier vector de longitud uno, puede ser tomado como primer vector de alguna base unitaria normal.

Una transformación T del espacio es *unitaria* si conserva las longitudes $|\xi T| = |\xi|$. Como en el Teorema 5, se deduce que una transformación unitaria conserva los productos internos y la ortogonalidad. Las ecuaciones $Y = XU$ representan una transformación unitaria si, y sólo si, la matriz U satisface a la condición $U^{*\prime}U = I$, donde U^* se deduce de U cambiando cada elemento de ésta por su conjugado (comparar con Teorema 7). Las matrices U con esta propiedad se llaman *unitarias;* como en el Teorema 6, cualquier fila de U tiene longitud uno, y dos filas de U son ortogonales. El conjunto de todas las transformaciones (o matrices) unitarias forma un grupo.

Las formas cuadráticas son ahora reemplazadas por formas «hermíticas», de las cuales el ejemplo más simple es la fórmula $\Sigma x_i x_i^*$ para la longitud.

En general, una *forma hermítica* es una expresión

$$(46) \qquad \sum x_i h_{ij} x_j^* = XHX^{*\prime}, \qquad H = \| h_{ij} \|$$

cuya matriz de coeficientes H tiene la propiedad $H^{*\prime}=H$. Una matriz H con esta propiedad se llama *hermítica;* en el caso especial de ser los términos h_{ij} reales, la matriz hermítica es simétrica. La forma (46) puede ser considerada como una función $h(\xi)=$ $=XHX^{*\prime}$ del vector ξ de coordenadas $x_1, \ldots, x_n$ relativas a cierta base. El valor $XHX^{*\prime}$ de esta función es siempre un número *real*. Para demostrarlo, bastará ver que tal número es igual a su conjugado (o, también, a su conjugado transpuesto). Pero como H es hermítica,

$$(XHX^{*\prime})^{*\prime}=(X^*H^*X^{**\prime})'=X'H^{*\prime}X^{*\prime}=XHX^{*\prime},$$

como decíamos.

Una transformación unitaria $Y=XU$, $X=YU^{-1}=YU^{*\prime}$ aplicada a una forma hermítica dará

$$XHX^{*\prime}=(YU^{-1})H(YU^{*\prime})^{*\prime}=YU^{-1}H(UY^{*\prime})=Y(U^{-1}HU)Y^{*\prime}.$$

La matriz de los coeficientes $U^{-1}HU$ es también hermítica,

$$(U^{-1}HU)^{*\prime}=U^{*\prime}H^{*\prime}(U^{-1})^{*\prime}=U^{-1}HU, \text{ ya que } U^{-1}=U^{*\prime}.$$

Se produce exactamente el mismo efecto sobre la forma cuando la referimos a un nuevo sistema unitario normal de coordenadas, ya que dará para ξ unas nuevas coordenadas Y relacionadas con las antiguas X por una ecuación $Y=XU$ con matriz U *unitaria*.

Interpretando de este modo la sustitución, se puede referir una forma hermítica a sus ejes principales. Los nuevos ejes serán escogidos por sucesivas propiedades de máximo, exactamente como en la reducción de una forma cuadrática real a sus ejes principales mediante transformaciones ortogonales. El primer eje α_1 se elige como el vector de longitud uno que hace máximo a $h(\xi)$; se puede entonces hallar una base unitaria normal a la que pertenezca α_1, por el Teorema 22. Para esta base desaparecen todos los términos rectangulares $x_1x_j^*$ para $j\neq 1$. Como los valores de la forma son siempre reales, los sucesivos máximos λ_i son asimismo números reales. Este proceso demuestra el siguiente teorema de los ejes principales :

TEOREMA 23. *Cualquier forma hermítica* $XHX^{*\prime}$ *puede reducirse a una forma diagonal real,*

(47) $$YHY^{*\prime}=\lambda_1y_1y_1^*+\lambda_2y_2y_2^*+\ldots+\lambda_ny_ny_n^*$$

por una transformación unitaria $Y=XU$.

Este teorema puede traducirse en otro relativo a la matriz H de una forma dada, como sigue :

TEOREMA 24. *Para cada matriz hermítica* H *existe una matriz unitaria* U *tal, que* $U^{-1}HU=U^{*\prime}HU$ *es una matriz diagonal real.*

Los métodos del Cap. X demostrarán también que los coeficientes diagonales λ_i de (47) son únicos.

EJERCICIOS

1. ¿Cuáles de las siguientes matrices son unitarias o hermíticas?

$$\begin{pmatrix} (1+i)/2 & (1-i)/2 \\ (1-i)/2 & (1+i)/2 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1-i \\ 1+i & \sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

2. Hallar una base unitaria y normal para el espacio de vectores ortogonales a $(1/2, i/2, (1+i)/2)$.
3. Demostrar que $\| h_{ij} \|$ es hermítica si, y sólo si, $h^*_{ij}=h_{ji}$ para todo i y j.
4. Demostrar que cualquier matriz A de números complejos puede escribirse en forma única como $A=H+S$, donde H es hermítica y S hemihermítica $(S^{*\prime}=-S)$.
5. Si G es una matriz de números complejos tal, que la expresión $XGX^{*\prime}$ es siempre real para cualquier complejo X, demostrar que G es hermítica. (Esto es la recíproca de una propiedad de las matrices hermíticas demostrada en el texto.)
6. Demostrar que todas las matrices $n \times n$ unitarias forman un grupo (llamado *grupo unitario*).
7. Demostrar el carácter lineal, hemisimétrico y positivo del producto interno (ξ, η).
8. Demostrar con detalle el Teorema 22, sobre base unitaria y normal.
9. a) Establecer con detalle el análogo de Teor. 5 para transformaciones unitarias.

   b) Lo mismo para Teor. 6.

   c) Lo mismo para Teor. 7.

*10. Demostrar un lema semejante al del § 9 para una forma hermítica en dos variables con un máximo en $x=0$, $y=1$. (*Sugerencia:* Dividir cada variable en sus partes real e imaginaria.)

*11. Dar una demostración detallada del teorema de los ejes principales para formas hermíticas.

12. Reducir la forma xy^*+x^*y a sus ejes principales.

13. Sean D y D_1 diagonales, y U y U_1 matrices unitarias. Demostrar que existen una matriz diagonal D_2 y una matriz unitaria U_2 tales, que $U^{-1}DU+U_1^{-1}D_1U_1=U_2^{-1}D_2U_2$.

* 12. Funciones y figuras

Desde un punto de vista geométrico, una función cuadrática $f(\xi)=XAX'+BX'+c$ se utiliza para definir una *figura* o *lugar*, que se obtiene igualando la función a cero. Esta figura está constituída por el conjunto de aquellos vectores ξ que satisfacen a la ecuación $f(\xi)=0$. En el espacio de dos dimensiones, la figura obtenida es, simplemente, una sección cónica. En el espacio tridimensional resulta una superficie cuádrica, y en general se obtendrá una *hipercuádrica* (o hipersuperficie cuádrica, como se dice también). Una transformación afín $Y=XP+K$ aplicada a la ecuación de esta hipersuperficie equivale, sencillamente, a efectuar la misma transformación sobre los puntos de la figura, obteniendo así una nueva figura, imagen o transformada de la original.

Es evidente que los resultados obtenidos en § 10 para la clasificación de funciones cuadráticas equivalentes darán una clasificacion similar para las correspondientes figuras geométricas. Observemos primero que una ecuación $f(\xi)=0$ y su producto por un escalar $cf(\xi)=0$ dan idéntica figura. Esto permite simplificar las formas canónicas obtenidas en § 10, tales como $y_1^2-y_2^2+c=0$. Si $c\neq 0$, esta ecuación representa el mismo lugar que la $(c^{-1})y_1^2-(c^{-1})y_2^2+1=0$; cuando $c>0$, ésta puede reducirse por la transformación afín $y_1=\sqrt{c}\,z_1$, $y_2=\sqrt{c}\,z_2$, a la forma $z_1^2-z_2^2+1=0$, mientras que si $c<0$, la transformación $y_i=\sqrt{-c}\,z_i$ da el resultado análogo $z_2^2-z_1^2+1=0$. En general, este proceso se puede aplicar para sustituir, por 1 o por 0, la constante c que aparece en (40) de § 10. Por lo tanto, en un espacio n dimensional afín, sobre el campo de los números reales, cualquier hipercuádrica es equivalente en el grupo afín a la figura determinada por alguna de las ecuaciones de los siguientes tipos:

$$y_1^2+\dots+y_p^2-y_{p+1}^2-\dots-y_r^2+y_{r+1}=0 \tag{48}$$

$$y_1^2+\dots+y_p^2-y_{p+1}^2-\dots-y_r^2+1=0 \tag{49}$$

$$y_1^2+\dots+y_p^2-y_{p+1}^2-\dots-y_r^2=0 \tag{50}$$

donde $0\leqslant p\leqslant r\leqslant n$, con $r<n$ en el caso (48).

Por ejemplo, en el plano, los posibles tipos de figuras con $r > 0$ son

$r=2$		$r=1$	
$x^2+y^2+1=0$	lugar vacío	$x^2+y=0$	parábola
$x^2-y^2+1=0$	hipérbola	$-x^2+y=0$	parábola
$-x^2-y^2+1=0$	círculo	$x^2+1=0$	lugar vacío
$x^2+y^2=0$	un punto	$-x^2+1=0$	dos rectas paralelas
$x^2-y^2=0$	dos rectas incidentes	$x^2=0$	una recta (doble)
$-x^2-y^2=0$	un punto.		

Obsérvese en particular que dos formas canónicas *diferentes*, x^2+y^2+1 y x^2+1, pueden dar *la misma* figura (esto es, la figura que no tiene ningún punto real).

La noción de equivalencia afín en las superficies cuádricas es también un ejemplo de la importante noción de equivalencia de figuras en un grupo de transformaciones. Las equivalencias en el grupo euclídeo se llaman congruencias; por ejemplo, dos triángulos del plano son congruentes si hay un movimiento rígido que lleve el primero sobre el segundo. De igual modo, dos triángulos son *semejantes* si existen una transformación por semejanza y una traslación que los lleve a coincidir. De modo que la relación de semejanza es exactamente una equivalencia en el grupo equiforme. También dos líneas rectas s y t son paralelas si existe una traslación que lleve s sobre t. Estos ejemplos pueden considerarse como casos particulares de la siguiente definición única:

DEFINICIÓN. *Sea* G *un grupo de transformaciones de un espacio* S. *Dos figuras de* S *son equivalentes en* G *si, y sólo si, existe en* G *alguna transformación* T *que a la primera figura haga corresponder la segunda.*

En esta definición, una *figura* f significa, simplemente, un conjunto de puntos P del espacio, y la transformación T lleva a f sobre la figura equivalente $f'=fT$, que consiste en todos los puntos PT, con P en f. La relación «f es equivalente a f'» es reflexiva, simétrica y transitiva: reflexiva, porque la transformación idéntica del grupo G representa a cada f sobre sí misma; simétrica, pues si T transforma f en f', T^{-1} pertenece a G y transforma f' en f; transi-

tiva, porque $fT=f'$, $f'U=f''$ implica que la transformación producto dé $fTU=f''$. Por lo tanto (Cap. VI, § 14), la relación de equivalencia en G proporciona una clasificación del conjunto de todas las figuras.

La anterior definición de equivalencia en un grupo se aplica a cualquier espacio. El *espacio* puede ser el conjunto de todos los enteros x ; el *grupo*, el conjunto de todas las traslaciones $x \to x+n$; la *figura*, un par de números enteros. Dos *figuras* (a, b) y (c, d) serán equivalentes en este grupo si, y sólo si, $b-a=d-c$. O bien, el espacio puede ser la superficie de una esfera, y el grupo, el conjunto de todas las rotaciones de una esfera sobre sí misma ; en tal caso, la «equivalencia» de triángulos esféricos coincide con la «congruencia» ordinaria. Otros ejemplos geométricos de equivalencia aparecerán en § 14.

EJERCICIOS

1. a) Enumerar todos los tipos posibles de superficies cuadráticas en el espacio tridimensional.
   b) Dar una breve descripción geométrica de cada tipo.
2. Clasificar: a) elipses, b) parábolas, y c) hipérbolas bajo el grupo equiforme (§ 2, al final). Hallar en cada caso un sistema completo de invariantes numéricos.
3. En un espacio vectorial $V_n(F)$ llamaremos *cuña* a la figura formada por dos vectores linealmente independientes β y γ. Demostrar que dos cuñas cualesquiera son equivalentes para el grupo lineal.
4. Demostrar que dos subespacios vectoriales son equivalentes bajo el grupo lineal si, y sólo si, tienen la misma dimensión.
5. Clasificar las superficies cuadráticas de un espacio euclídeo n-dimensional, para el grupo de los movimientos rígidos (utilizar el Teorema 20).

* 13. Subespacios afines

En el espacio de los vectores reales de dos dimensiones $V_2(R^*)$, los subespacios vectoriales propios son rectas que pasan por el origen. Cualquier recta L que no pase por el origen, puede ser trasladada a otra recta que pase por él, de modo que L es equivalente en el grupo de traslaciones a un subespacio lineal del espacio vectorial. Esto sugiere el estudio general de los conjuntos equivalentes en las afinidades de los subespacios lineales que pasan por el origen. Tales conjuntos son llamados subespacios afines. Así, un *subespacio afín* M en un espacio vectorial $V_n(F)$ es un conjunto que resulta de un subespacio lineal S_0 por una transformación afín

$\xi \to \xi T + \lambda$. Sabemos que la transformación lineal T solamente cambia el subespacio lineal S_0 en otro subespacio lineal S; luego cualquier subespacio afín M puede ser obtenido por una traslación $\xi \to \xi + \lambda$ aplicada a un subespacio lineal S. Ahora bien, $M = S + \lambda$ consistirá en todos los vectores trasladados $\xi + \lambda$ para ξ en S. Diremos que $S + \lambda$ es «paralelo» a S y su dimensión es, por definición, la dimensión de S. La diferencia de dos vectores de M está en S, luego M determina unívocamente un subespacio lineal paralelo S. La traslación por λ, sin embargo, no está determinada de modo único: si $\mu = \alpha + \lambda$ es cualquier vector de $S + \lambda$, entonces μ y λ dan el mismo subespacio afín $S + \mu = S + \lambda$. Como cualquiera de estos subespacios se obtiene sumando un elemento fijo al subgrupo aditivo S del grupo V, el suespacio $S + \lambda$ puede definirse como la clase de restos según S (Cap. VI, § 9) que contiene a λ.

Una recta es un subespacio afín $S_1 + \lambda$ de una dimensión, obtenida por traslación del subespacio vectorial S_1 de una dimensión. Si S_1 tiene base ϵ, consiste en todos los productos $t\epsilon$ de ϵ por escalares t; así que la recta $S_1 + \lambda$ consiste en todos los vectores $t\epsilon + \lambda$ (ϵ y λ vectores fijos, t escalar variable).

Dos vectores distintos α y β (esto es, sus extremos) están sobre una recta única. Pues en una recta $S_1 + \lambda$ puede tomarse para λ uno de los dos vectores, α por ejemplo. El subespacio paralelo S_1 debe contener a la diferencia $\beta - \alpha \neq 0$ como una base, así que $S_1 + \alpha$ consiste en todos los vectores $\xi = \alpha + t(\beta - \alpha)$. Si ξ, α, β tienen coordenadas x_i, a_i, b_i, estas ecuaciones pueden escribirse $x_i = a_i + t(b_i - a_i)$, $i = 1, \ldots, n$. Estas son las ecuaciones «paramétricas» de la recta que une los dos puntos dados. Si $n = 2$ y $b_i \neq a_i$, resulta

$$t = (x_1 - a_1)/(b_1 - a_1) = (x_2 - a_2)/(b_2 - a_2),$$

que es la conocida fórmula que da la geometría analítica para la recta que une los puntos (a_1, a_2) y (b_1, b_2).

Un plano ordinario se caracteriza frecuentemente por la siguiente propiedad: si contiene a dos puntos de una recta cualquiera, contiene a todos los puntos de ella. La misma propiedad es válida para un subespacio afín de *cualquier* dimensión sobre *cualquier* campo. Porque, dados dos vectores $\alpha = \alpha_0 + \lambda$, $\beta = \beta_0 + \lambda$ en un subespacio $S + \lambda$, la recta de los vectores

$$\xi = \alpha + t(\beta - \alpha) = \alpha_0 + \lambda + t(\beta_0 - \alpha_0)$$

pertenece evidentemente a $S+\lambda$ para todo t. Vamos a demostrar que si en el campo de las coordenadas es $1+1\neq 0$, la recíproca es cierta, esto es,

TEOREMA 25. *Un subconjunto* M *no vacío del espacio* $V_n(F)$ *es un subespacio afín si, y sólo si, la recta que une dos puntos de* M *está contenida por completo en* M *(supuesto que* $1+1\neq 0$ *en* F).

Demostración. Dado M con la propiedad enunciada, nos proponemos, ante todo, construir un subespacio vectorial «paralelo» que pase por el origen. Escogido un vector λ de M, traslademos según λ, y denotemos por $S=M-\lambda$ el conjunto resultante de todas las diferencias $\eta-\lambda$, con η en M. Para probar que $M=S+\lambda$ es un subespacio afín, bastará demostrar que S es un subespacio vectorial. Como las rectas se trasladan a otras rectas, la hipótesis sobre M asegura la propiedad análoga para S: la recta que une dos vectores de S yace en S. Para cualquier α en S, la recta que une 0 (en S) con α yace en S, el cual contiene, por lo tanto, a todos sus múltiples escalares $c\alpha$. Si S contiene a α y a β, contendrá a 2β y 2α, luego también a toda la recta $\xi=2\alpha+t(2\beta-2\alpha)$, que ellos determinan (¡hágase la figura!). En particular, para $t=1/2$ contendrá $\xi=2\alpha+(\beta-\alpha)=\beta+\alpha$, que es la suma de los dos vectores dados. Hemos demostrado así que S es cerrado para la suma y el producto por escalares, luego es un subespacio vectorial, como queríamos demostrar. Obsérvese, sin embargo, que $1+1\neq 0$ es condición necesaria porque interviene el coeficiente 1/2. Así, el teorema no es válido para una «geometría» afín sobre el campo de los enteros módulo 2.

En el plano real, el punto medio (x, y) del segmento que une (x_1, y_1) con (x_2, y_2) está dado por $x=(x_1+x_2)/2$, $y=(y_1+y_2)/2$. Por analogía, el «punto medio» de los vectores ξ_1 y ξ_2 se define como $\xi_3=(1/2)(\xi_1+\xi_2)$, suponiendo que $1+1\neq 0$ en el campo de escalares. En una transformación afín, el punto medio de los transformados η_1, η_2 de ξ_1, ξ_2 será:

$$(\eta_1+\eta_2)/2=(\xi_1T+\kappa+\xi_2T+\kappa)/2=[(\xi_1+\xi_2)/2]T+\kappa.$$

Por lo tanto, el nuevo punto medio resulta ser el transformado del punto medio original. Esto demuestra

TEOREMA 26. *Una transformación afín de un espacio* $V_n(F)$ *hace corresponder los puntos medios con los puntos medios (supuesto* $1+1 \neq 0$ *en* F).

En el caso particular del espacio vectorial euclídeo, puede demostrarse mucho más. El segmento rectilíneo de extremos α, β se llamará paralelo al de extremos γ, δ cuando para algún escalar c no nulo se verifica la igualdad $(\delta-\gamma)=c(\beta-\alpha)$. El escalar c se llama la *razón* entre los dos segmentos, ya que coincide con la razón ordinaria para segmentos de un plano, pues para las longitudes de estos vectores se encuentra que $|\delta-\gamma|=|c||\beta-\alpha|$. Además, en una transformación afín (1) que lleve α a α', β a β', δ a δ' y γ a γ, un cálculo directo muestra que $(\delta'-\gamma')=(\delta-\gamma)T=c(\beta-\alpha)T=$ $=c(\beta'-\alpha')$. Esto demuestra

TEOREMA 27. *Una transformación afín de un espacio vectorial euclídeo conserva el paralelismo entre segmentos rectilíneos y deja invariante la razón entre las longitudes de los mismos.*

Inversamente, este resultado significa que todas las longitudes paralelas a una dada quedan multiplicadas por la misma constante. Naturalmente, las distancias en direcciones diversas pueden quedar multiplicadas por constantes distintas, como se ve si pensamos que las afinidades incluyen corrimientos, compresiones y dilataciones.

Apéndice. La propiedad del Teorema 27 puede ser utilizada para definir directamente la transformación afín [sin intermedio de las transformaciones lineales, como en nuestra definición (1)].

TEOREMA 28. *Una transformación* $\xi \to \xi H$ *de un espacio vectorial es afín si, y sólo si, cada par de segmentos rectilíneos paralelos se transforma en otro par de segmentos paralelos y la razón entre las longitudes de los segmentos de cada par es la misma.*

Demostración. Si H tiene la propiedad enunciada, $\beta-\alpha=$ $=c(\delta-\gamma)$ implicará $\beta H-\alpha H=c(\delta H-\gamma H)$. Utilizaremos el transformado OH del vector cero para definir una nueva operación T por $\xi T=\xi H-OH$. Como aquí OH representa una traslación, bastará probar que T es lineal. Primero consideremos el producto por un escalar, $c\xi$. La relación $c\xi-O=c(\xi-O)$ implica por la hipótesis que $(c\xi)H-OH=c(\xi H-OH)$, o que $(c\xi)T=c(\xi T)$. Para el

caso de una suma $\xi+\eta$, observemos que el paralelogramo de ξ y η tiene como lados opuestos los segmentos $(\xi+\eta,\ \xi)$ y $(\eta,\ O)$, paralelos y de razón 1, pues $(\xi+\eta)-\xi=1\cdot(\eta-O)$. Por hipótesis, H conserva esta razón, luego

$$(\xi+\eta)H-\xi H=1(\eta H-OH), \qquad (\xi+\eta)T-\xi T=\eta T.$$

Estas dos conclusiones prueban que T es lineal, c. q. d.

EJERCICIOS

1. Para cada uno de los siguientes pares de puntos, hallar la ecuación paramétrica de la recta que les une y representarla en la forma $S_1+\lambda$ (esto es, hallar el espacio S_1).
   a) (2, 1) y (5, 0);
   b) (4, —2) y (2, —6);
   c) (1, 3, 2) y (—1, 7, 5);
   d) (4, 5, —1) y (0, 5, 5);
   e) (1, 2, 3, 4) y (4, 3, 2, 1).
2. Representar la línea entre (1, 3) y (4, 2) en la forma $S+\lambda$, con cuatro valores distintos de λ. Dibujar una figura.
3. Demostrar: por tres vectores α, β, γ (esto es, por sus extremos) no alineados, pasa un solo espacio bidimensional afín (¡un plano!). Demostrar que los vectores de este plano tienen la forma $\xi=\alpha+s(\beta-\alpha)+t(\gamma-\alpha)$, con s y t escalares variables.
4. Hallar las ecuaciones paramétricas (en la forma de Ejercicio 3) para el plano (si existe) determinado por cada una de las siguientes ternas de puntos:
   a) (1, 3, 2), (4, 1, —1), (2, 0, 0);
   b) (1, 1, 0), (1, 0, 1), (0, 1, 1);
   c) (2, —1, 3), (1, 1, 1), (3, 0, 4).
5. En cada parte del Ejercicio 4, hallar una base para el subespacio lineal paralelo que pasa por el origen.
6. Enunciar y demostrar un teorema análogo al Ejercicio 3, para subespacios de dimensión cualquiera.
7. Demostrar que el Teorema 25 es falso para el «espacio» vectorial de dos dimensiones sobre el campo de enteros mód. 2.

*8. Demostrar que en una transformación afín de $V_2(F)$ que deje fijos tres puntos no alineados, deben quedar fijos todos los puntos del espacio.

9. Por definición, un hiperplano en $V_n(F)$ es un subespacio afín de dimensión $n-1$.
   a) Demostrar que el conjunto de todos los vectores ξ cuyas coordenadas satisfacen a una ecuación lineal $\sum_i a_i x_i = c$ es un hiperplano, supuesto que los coeficientes a_i no son todos nulos.
   b) Probar que, recíprocamente, cualquier hiperplano tiene tal ecuación.
   c) Hallar la ecuación del hiperplano que contiene a (1, 0, 1, 0), (0, 1, 0, 1), (0, 1, 1, 0), (1, 0, 0. 1).

10. Demostrar que cualquier subespacio afín puede representarse por ecuaciones lineales no homogéneas, análogamente al Ejercicio 9.
11. Sea V un espacio vectorial sobre un campo ordenado F. Llamaremos «centroide» de los puntos $\xi_1, \ldots, \xi_r$, con los respectivos «pesos» $m_1, \ldots, m_r$, al punto $\xi' = (m_1\xi_1 + \ldots + m_r\xi_r)/M$, donde $M = m_1 + \ldots + m_r$. Demostrar: a) cualquier transformación afín transporta al centroide sobre el nuevo centroide; b) cualquier transformación que conserva los centroides es afín.

14. Otras aplicaciones geométricas

La equivalencia en el grupo afín (que según parece no fué conocida por Euclides) tiene gran número de interesantes aplicaciones elementales. En el grupo afín, dos triángulos cualesquiera son equivalentes. Para demostrarlo, bastará probar que un triángulo cualquiera $\alpha\beta\gamma$ es equivalente al triángulo *equilátero* de vértices $O = (0, 0)$, $\beta_0 = (2, 0)$ y $\gamma_0 = (1, \sqrt{3})$ (ver fig. 2). Por una traslación, el vértice α puede llevarse al origen O, los otros dos vértices ocuparán las posiciones β' y γ'. Como estos vectores β' y γ' son linealmente independientes, existirá una transformación lineal $x\beta' + y\gamma' \to x\beta_0 + y\gamma_0$ que lleve β' a β_0, γ' a γ_0. El producto de la traslación por esta transformación lineal llevará a $\alpha\beta\gamma$ sobre $O\beta_0\gamma_0$, como deseábamos.

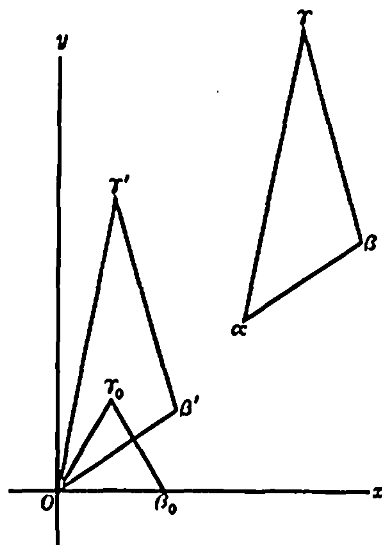


Figura 2

Así, cualquier triángulo es equivalente a uno equilátero. Pero en este último, las tres medianas, por simetría, deben cortarse en un mismo punto (centro de gravedad). Una transformación afín transforma los puntos medios en puntos medios. Esto demuestra la propiedad elemental de que las tres medianas de un triángulo cualquiera se cortan en un punto. Además, se ve en seguida que el punto de intersección de las medianas en un triángulo equilátero las divide en la relación $1 : 2$; luego la misma propiedad vale para cualquier triángulo (Teor. 27).

Por otra parte, una elipse es equivalente afín a un círculo. Pero cualquier diámetro pasa por el centro del círculo, y tiene tangentes paralelas en sus extremos; y además, el *diámetro conjugado*, que es el paralelo a estas tangentes, biseca todas las cuerdas paralelas al diámetro dado. Se deduce de aquí que las mismas propiedades valen para cualquier elipse, pues una transformación afín conserva el paralelismo y hace corresponder tangentes con tangentes (pero se observará que los diámetros conjugados en una elipse no son ortogonales).

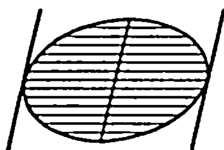


Figura 3

Las equivalencias pueden emplearse para sustituir unas figuras por otras más simples pero equivalentes. Así, cualquier triángulo es equivalente afín de otro equilátero de lado 1; cualquier elipse es equivalente afín al círculo de radio 1 con centro en el origen; cualquier vector $\xi \neq 0$ es linealmente equivalente al vector unidad $(1, 0, ..., 0)$, y así sucesivamente. Cuando esta forma sencilla equivalente está determinada de modo único, se le llama canónica. Más explícitamente, dado cualquier tipo de figuras, se dice que una selección de ellas es un conjunto de *formas canónicas*, si cada figura del tipo en cuestión es equivalente a una, y sólo a una, de estas formas canónicas. Este concepto es análogo al de la expresión canónica de una forma cuadrática en un grupo, tal como se trató en § 5.

Para los triángulos isósceles, por ejemplo, es evidente que cada uno es equivalente por el grupo euclídeo a uno, y sólo a uno, de los triángulos isósceles que tengan la base sobre el eje X y la altura sobre el eje Y (positivo). Dos triángulos distintos en esta posición particular no pueden ser equivalentes en el grupo, así que este conjunto de triángulos constituye un sistema canónico para el conjunto de todos los triángulos isósceles.

Los triángulos isósceles equivalentes pueden ser descritos por completo mediante dos «invariantes» numéricos, la longitud de la base y la de la altura. En general, un invariante tiene también aquí la misma significación que para una forma cuadrática (§ 5): Sea una regla (es decir, una función) que a cada figura de un tipo dado le asigne un número único $J(f)$. Si para cualquier transformación T de un grupo dado G es $J(f) = J(fT)$, la función $J(f)$ se llama un *invariante* de f *en* (o *para*) *el grupo* G. Por ejemplo, la

longitud de un segmento es invariable en el grupo euclídeo, ya que dos segmentos son equivalentes en este grupo si, y sólo si, tienen la misma longitud. En general, los invariantes $J_1, \ldots, J_r$ de una clase de figuras constituyen un *sistema completo de invariantes* si dos figuras f y f' de la clase son equivalentes cuando tengan los mismos invariantes, $J_1(f)=J_1(f'), \ldots, J_r(f)=J_r(f')$, y sólo en este caso. (Ejemplo: dos triángulos son equiformes si, y sólo si, sus ángulos son iguales a pares.)

EJERCICIOS

1. Demostrar que cualquier paralelogramo es equivalente afín de un cuadrado.
2. Dar una demostración afín de que las diagonales de un paralelogramo se bisecan.
3. a) Hallar una transformación afín que transforme el triángulo de vértices (0, 0), (0, 1) y (1, 0) en el triángulo equilátero con vértices (1, 0), (—1, 0), (0, $\sqrt{3}$).
   b) El mismo problema, si el primer triángulo tiene vértices (1, 1), (1, 2) y (3, 3).
4. Describir geométricamente una sucesión de traslaciones, rotaciones, corrimientos y compresiones, que transforme un triángulo de posición arbitraria en un triángulo equilátero dado.
5. Demostrar por métodos afines que en un trapecio las dos diagonales y la línea que une los puntos medios de los lados paralelos pasan por un mismo punto.
6. Demostrar que cualquier paralelepípedo es equivalente afín de un cubo.
7. Demostrar que las cuatro diagonales de cualquier paralelepípedo tienen el punto medio común. (Es el centro de gravedad.)
8. * Demostrar que en un elipsoide los centros de gravedad de las diversas familias de secciones planas paralelas están en una línea recta.
9. Hallar un sistema completo de invariantes para una elipse en las transformaciones de semejanza.
10. Hallar un sistema completo de invariantes de un triángulo para el grupo euclídeo.
11. Hallar las formas canónicas y un conjunto completo de invariantes para las figuras planas de cada uno de los siguientes tipos, tanto en el grupo euclídeo como en el grupo equiforme: a) triángulos rectángulos; b) paralelogramos; c) rectángulo.

CAPÍTULO X

Característica y determinante de una matriz

1. La característica y los sistemas homogéneos

En general, el *dominio* de una transformación (o función) T se define como el conjunto de los elementos originales, esto es, aquellos a quienes T se aplica, mientras que se llama *resultante* de T al conjunto de sus representaciones o «imágenes» (esto es, la imagen del dominio por T).

Cuando T es una transformación *lineal* homogénea de un espacio vectorial V sobre un nuevo espacio vectorial W, el resultante (conjunto de todos los ξT), no puede ser un subconjunto arbitrario de W.

LEMA 1. *Una transformación lineal de un espacio vectorial V da como resultante otro espacio vectorial (y por lo tanto, un subespacio de W).*

Demostración. Como $c(\xi T) = (c\xi)T$ y $\xi T + \eta T = (\xi + \eta)T$, el conjunto de elementos «imagen» es cerrado para las dos operaciones sobre los vectores.

La transformación T, referida a las respectivas bases del espacio m-dimensional V y del n-dimensional W, vendrá representada por una ecuación lineal $Y = XA$, como en Cap. VIII, (29). Se llamará *resultante* de la matriz A (que es $m \times n$) al resultante de esta transformación; en otras palabras, se llama resultante de A al conjunto de todos los vectores XA, para cualquier $X = (x_1, \ldots, x_m)$. La T transforma cada X en su correspondiente

$$Y = XA = (\Sigma x_i a_{i1}, \ldots, \Sigma x_i a_{in}) = \Sigma x_i (a_{i1}, \ldots, a_{in}).$$

así que el resultante de A consiste en todas las combinaciones lineales de las filas $A_i = (a_{i1}, \ldots, a_{in})$ de A. Es, por consiguiente, el subespacio de $V_n(F)$ engendrado por las filas de A.

DEFINICIÓN. *Se llama característica de una transformación lineal* T *o matriz* A, *al número de dimensiones de su resultante.*

Como el número de dimensiones de un subespacio es el máximo número de sus elementos linealmente independientes, la característica de A es también el número máximo de filas de A linealmente independientes. Por este motivo, la característica de A que ahora hemos definido se llama a veces *característica por filas*, para distinguirla de la *característica por columnas*, que es el número máximo de columnas de A linealmente independientes. Ésta coincide, desde luego, con la característica por filas de la transpuesta A' de A (a partir del Cor. 3 del Teor. 14 no será preciso mantener esta distinción).

Concepto dual al de resultante de una matriz o de una transformación lineal es el de su *espacio nulo.*

DEFINICIÓN. *El espacio nulo de una transformación lineal es el conjunto de todos los vectores* ξ *tales, que* ξT$=0$. *El espacio nulo de una matriz* A *es el conjunto de todas las matrices de una fila* X, *que satisfacen a la ecuación lineal homogénea* XA$=0$.

LEMA 2. *El espacio nulo de una transformación lineal (o matriz) es un subespacio de su dominio.*

Demostración. Si $\xi T = 0$ y $\eta T = 0$, también para todo c y c' será

$$(c\xi + c'\eta)T = c(\xi T) + c'(\eta T) = 0 + 0 = 0.$$

Por lo tanto, $c\xi + c'\eta$ pertenece al espacio nulo, el cual es, por lo tanto, un subespacio lineal.

La dimensión (lineal) del espacio nulo de una matriz A o de una transformación lineal T, se llama *nulidad* de A o T. Nulidad y característica están relacionadas por una igualdad fundamental, válida en las matrices y en las transformaciones lineales. A causa de la correspondencia entre ambos conceptos, bastará exponer la demostración en el caso de las matrices.

TEOREMA 1. *Característica + Nulidad = Dimensión del dominio.*

Así, para las matrices $m \times n$, la característica (por filas) más la nulidad (por filas) es igual a m. Y para una transformación lineal de $V_n(F)$ en sí mismo, característica + nulidad = n.

Demostración. Si la nulidad de T es s, el espacio nulo N tendrá una base $\alpha_1, \ldots, \alpha_s$ de s elementos, la que puede extenderse a una base $\alpha_1, \ldots, \alpha_s, \beta_1, \ldots, \beta_r$ para todo el dominio de T. Como cualquier $\alpha_i T = O$, los vectores $\beta_i T$ engendran la resultante R de T. Además, $x_1(\beta_1 T) + \ldots + x_r(\beta_r T) = O$ implica que $x_1\beta_1 + \ldots + x_r\beta_r$ pertenece a N, así que $x_1 = \ldots = x_r = 0$. Por lo tanto, los vectores $\beta_i T$ son independientes y forman una base de R. Resulta así que la dimensión $s + r$ del dominio es la suma de la dimensión s de N con la r de R, c. q. d.

TEOREMA 2. *Para que una matriz* m × n, A, *sea regular, es necesaria y suficiente cualquiera de las dos condiciones siguientes: a) Característica de* A = n ; *b) nulidad de* A = 0.

Demostración. Por el Teorema 1, a) y b) son equivalentes entre sí. Podemos, pues, limitarnos a demostrar a). Pero a) equivale a afirmar que el resultante de A es todo el espacio $V_n(F)$ de matrices de una fila. Esta condición, a su vez, equivale a la regularidad de A, como se ha demostrado en Cap. VIII, Teor. 7.

Las nociones de característica y nulidad son importantes en la teoría de las ecuaciones lineales homogéneas. Consideremos un sistema de tales ecuaciones

$$(1) \qquad \begin{array}{c} x_1 a_{11} + x_2 a_{21} + \ldots + x_n a_{n1} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_1 a_{1m} + x_2 a_{2m} + \ldots + x_n a_{nm} = 0. \end{array} \qquad \begin{array}{l} (m \text{ ecuaciones,} \\ n \text{ incógnitas)} \end{array}$$

Si representamos los coeficientes a_{ji} en orden transpuesto mediante la matriz $A = \| a_{ij} \|$, se tendrá para (1) la forma abreviada

$$(1') \qquad XA = O.$$

El espacio nulo de A es entonces exactamente el conjunto de todas las soluciones de (1), es decir : el conjunto de todas las soluciones (1) es un subespacio del espacio vectorial. De aquí se desprende inmediatamente un corolario.

TEOREMA 3. *Si la matriz transpuesta* A *de los coeficientes de* m *ecuaciones lineales homogéneas con* n *incógnitas tiene característica* r, *todas las soluciones podrán expresarse como combinaciones lineales de* n — r *soluciones linealmente independientes. Luego si* A *tiene nulidad* s = n — r, *el subespacio de las soluciones de* XA = O *tiene una base constituida por* s *soluciones.*

COROLARIO. n *ecuaciones lineales homogéneas con* n *incógnitas* x_i *tendrán una solución distinta de* $x_1 = \ldots = x_n = 0$ *si la matriz de los coeficientes es singular, y sólo en este caso.*

Las ecuaciones lineales homogéneas son susceptibles de una interpretación geométrica directa. Una sola ecuación

$$(2) \qquad x_1a_{11} + \ldots + x_na_{n1} = 0, \quad \text{alguna } a_{i1} \neq 0,$$

tiene una matriz de coeficientes $n \times 1$, de característica 1, así que, por el Teorema 3, las soluciones forman un subespacio $(n-1)$-dimensional, del espacio $V_n(F)$ de todos los vectores $(x_1, \ldots, x_n)$. Para el caso del plano, o del espacio ordinario, esto es un hecho muy sabido en la Geometría Analítica. En cualquier V_n, un subespacio lineal $(n-1)$-dimensional se llama *hiperplano.*

Por lo tanto, los vectores $(x_1, \ldots, x_n)$ que satisfacen a m ecuaciones simultáneas $XA_1 = \ldots = XA_m = 0$ determinan una figura que es la *intersección* de m hiperplanos $H_1, \ldots, H_m$ que pasan por el origen O. Para cualquier subespacio S, la $d[H_m + S]$ es $(n-1)$ o n, luego la identidad dimensional del Teor. 9 del Cap. VII puede escribirse así:

$$(3) \qquad d[H_m \frown S] = d[S] - d[H_m + S] + d[H_m] \geq d[S] - 1.$$

Por inducción sobre m obtenemos $d[H_1 \frown \ldots \frown H_m] \geq n - m$. Esto significa que m ecuaciones lineales homogéneas con n incógnitas tienen por lo menos $n - m$ soluciones linealmente independientes.

EJERCICIOS

1. Hallar las resultantes, espacios nulos, características y nulidades de las transformaciones dadas en § 1, Cap. VIII, Ejercs. 1 a) - 1 d) y 4 a) - 4 b).
2. Construir una transformación de V_3 sobre sí mismo cuya resultante esté engendrada por los vectores (1, 3, 2) y (3, —1, 1).
3. Construir una transformación de V_4 sobre sí mismo que tenga como espacio nulo el engendrado por (1, 2, 3, 4) y (2, 2, 4, 4).

4. Hallar una base para el conjunto de las soluciones linealmente independientes de cada uno de los cuatro sistemas de ecuaciones que siguen:
   a) $x+y+3z=0$, $2x+2y+6z=0$;
   b) $x+y+z=0$, $y+z+t=0$;
   c) $x+2y-4z=0$, $3x+y-2z=0$;
   d) $x+y+z+t=0$, $2x+3y-z+t=0$, $3x+4y+2t=0$.
5. Resolver el Ejerc. 4, si las igualdades se hacen congruencias mód. 5.
6. Demostrar que las características por filas de un producto AB no exceden a la característica por filas de B.
7. Si la matriz $n\times n$ A es regular, demostrar que, para cualquier matriz $n\times n$ B, las matrices AB, B y BA tienen la misma característica.
8. Demostrar que característica $(A+B)\leqslant$ característica (A) + característica (B).
9. Dadas las características de A y B, ¿cuál es la característica de $\begin{pmatrix}0 & B\\ A & 0\end{pmatrix}$?
10. Completar la demostración por inducción de la expresión que sigue a (3). ¿Por qué es (3) una desigualdad y no una igualdad?
11. Si T es una transformación lineal de V_n con resultante R, mientras que S es cualquier subespacio de V_n que contiene a R, demostrar que T da origen a una transformación lineal T^* de S. Demostrar que la característica de T^* no supera a la de T.

. Matrices equivalentes por filas

La técnica para resolver sistemas de ecuaciones lineales descrita n Cap. II, § 3, puede ser parafraseada para conseguir un método e calcular la característica de cualquier matriz $m\times n$ sobre un camo arbitrario F. Si A es regular, esta paráfrasis da también un método para calcular la inversa de A.

El método práctico se reduce al uso reiterado de tres tipos de *peraciones elementales con las filas* de una matriz A:

a) El intercambio de dos filas cualesquiera.
b) La multiplicación de una fila por un escalar $c\neq 0$ de F.
c) La adición de una fila con otra (*).

La operación de sumar a una fila A_j el producto de la fila i or el escalar c, es una combinación de estas «operaciones elemenales»; primero se multiplica la fila i por c, luego se suma esta ueva fila i a la fila j, finalmente se multiplica la nueva fila i por c^{-1}.

(*) Dicho con más exactitud: c) Reemplazar la fila A_j de A por la suma A_j+A_i e las filas j-ésima e i-ésima, con $i\neq j$.

DEFINICIÓN. *Dos matrices* m × n, A *y* B, *se llaman equiva ntes por filas si* B *puede obtenerse a partir de* A *por una succ- ión de operaciones elementales entre sus filas.*

El efecto de cada operación elemental puede deshacerse por tra operación del mismo tipo, así que la relación de equivalencia or filas es simétrica (si B equivale a A, A equivale a B). El ca- ácter reflexivo y el transitivo son patentes.

Con operaciones sobre filas se puede siempre reducir una ma ;iz a una forma triangular. Dada A, elijamos, si existe, alguna la en la que el primer elemento b no sea nulo, y permutemos ıs filas para colocarla en primer lugar. Multiplicando la primera la por b^{-1} tendremos el primer elemento reducido a 1. Todos los tros elementos de la primera columna pueden ser reducidos a cero, ımando a su fila un múltiplo conveniente de la primera, así que ι matriz toma una de estas dos formas:

$B = (0, C)$, con C $m \times (n-1)$ y 0 $m \times 1$, o bien

$$B = \begin{pmatrix} 1 & Z \\ 0 & C \end{pmatrix}, \text{ con}$$

$$\begin{Bmatrix} B & Z & 0 & C \\ m \times n & 1 \times (n-1) & (m-1) \times 1 & (m-1) \times (n-1) \end{Bmatrix}.$$

Repitamos este proceso para el bloque C. Tomaremos como pri- ıera fila de C la que tenga el primer elemento no nulo (si la hay), educiremos a 1 este elemento y a 0 los restantes elementos de la rimera columna de C. El resultado final será una matriz con los eros repartidos, como, por ejemplo, en la siguiente matriz 4×7:

$$\begin{bmatrix} 0 & 1 & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ 0 & 0 & 0 & 1 & d_{25} & d_{26} & d_{27} \\ 0 & 0 & 0 & 0 & 1 & d_{36} & d_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

En cada fila, el primer elemento no nulo que interviene es un 1 el número de ceros que le precede es mayor en cada fila que en ι precedente. Una matriz de esta forma se llama *escalonada*. Con › anterior hemos demostrado

TEOREMA 4. *Cualquier matriz es equivalente por filas a una ıatriz escalonada.*

La característica de una matriz escalonada D puede leerse con una mirada. Supongamos que D tiene r filas $D_1, \ldots, D_r$ no totalmente nulas y que hubiese alguna relación lineal $c_1D_1 + \ldots + c_rD_r = 0$ entre ellas. En esta relación no podría intervenir D_1, porque su primera coordenada es 1 y en las restantes filas es cero. Luego $c_1 = 0$. La relación reducida $c_2D_2 + \ldots + c_rD_r = 0$ no podría tener a D_2 por la misma razón; por lo tanto, $c_1 = \ldots = c_r = 0$, y las filas son independientes. Igualmente se ve la independencia lineal entre las r columnas, cada una de las cuales contiene un elemento 1 de los ahora considerados; y como cada columna tiene a lo más r elementos no nulos, las columnas son, en esencia, vectores de un espacio r-dimensional, así que no puede haber más de r columnas linealmente independientes (Teor. 7, Cor. 2, Cap. VII). Así hemos demostrado

TEOREMA 5. *Una matriz escalonada con* r *filas no nulas tiene característica (por filas y por columnas) igual a* r.

Demostraremos pronto (Teoremas 10 y 14) que dos matrices equivalentes por filas tienen la misma característica por filas y la misma característica por columnas. Por lo tanto, *para computar la característica* (y la característica por columnas) *de cualquier matriz* A, *basta formar la correspondiente matriz escalonada* D (como en el Teor. 4) *y contar el número de las filas no nulas de* D.

Para probar que dos matrices equivalentes por filas tienen la misma característica, es conveniente interpretar las operaciones elementales entre filas como premultiplicaciones por factores convenientes. Por ejemplo, dos filas de una matriz pueden permutarse multiplicando la matriz dada por la que se obtiene permutando las filas de la matriz idéntica I, como

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = \begin{pmatrix} 0\cdot a_1 + 1\cdot b_1 & 0\cdot a_2 + 1\cdot b_2 \\ 1\cdot a_1 + 0\cdot b_1 & 1\cdot a_2 + 0\cdot b_2 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 \\ a_1 & a_2 \end{pmatrix}.$$

Para sumar la segunda fila a la primera, o para multiplicar la segunda fila por c, hagamos lo mismo con el factor identidad de la izquierda:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 & a_2 + b_2 \\ b_1 & b_2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ cb_1 & cb_2 \end{pmatrix}.$$

Resultados análogos son válidos para las matrices $n \times n$; los prefactores empleados para representar estas operaciones se llaman *matrices elementales*.

DEFINICIÓN. *Una matriz elemental* $n \times n$, E, *es cualquiera de las matrices que resultan de la matriz idéntica* I *aplicándole una operación elemental entre filas.*

Tenemos así tres tipos de matrices elementales, que pueden describirse como sigue: Designemos por I_k la fila k-ésima de la matriz idéntica I. Con esto, el intercambio en I de las filas de orden i y j dará una matriz elemental $H = H_{ij}$ cuyas filas H_k serán

$$(4) \qquad H_i = I_j, \quad H_j = I_i, \quad H_k = I_k \quad (k \neq i, j).$$

Ejemplos de matrices elementales son

$$\underset{H_{24}}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}, \quad \underset{I+2E_{33}}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}, \quad \underset{F_{12}}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}.$$

Análogamente, al multiplicar en I la fila i por un escalar c, resultará la matriz M cuyas filas M_k serán

$$(5) \qquad M_i = cI_i, \quad M_k = I_k \quad (k \neq i).$$

Si E_{ii} es la matriz que tiene un elemento 1 en el cruce de la fila i con columna i, y los restantes elementos nulos, esta matriz M puede escribirse así: $M = I + (c-1)E_{ii}$. Finalmente, la operación elemental de adicionar la fila i a la fila j, aplicada a I, dará la matriz elemental F_{ij} cuyas filas F_k están dadas por

$$(6) \qquad F_j = I_i + I_j, \quad F_k = I_k \quad (k \neq j).$$

TEOREMA 6. *Cada operación elemental entre las filas de una matriz* $m \times n$, A, *es resultado de una premultiplicación por la correspondiente matriz elemental* E.

Esto puede probarse fácilmente por cálculo directo del producto EA, como en el precedente caso 2×2. Consideremos, por ejem-

plo, la operación elemental de adicionar las filas i y j de A. Las filas F_k de la correspondiente matriz elemental $F_{ij}=F$ están dadas por (6). Las filas de cualquier producto EA se obtienen a partir de las filas del primer factor, por la fórmula (33) del Cap. VIII:

$$(FA)_j = F_jA = (I_i + I_j)A = I_iA + I_jA = (IA)_i + (IA)_j,$$
$$(FA)_k = F_kA = I_kA = (IA)_k \qquad (k \neq j).$$

Estas igualdades muestran que las filas de F se obtienen de las de $IA=A$ por adición de la fila i-ésima a la fila j-ésima. Es decir, esta operación elemental transforma A en FA, como asegura el Teorema 6.

COROLARIO 1. *Una matriz elemental E es regular.*

Demostración. E se obtiene de I por cierta operación. La operación inversa corresponde a alguna matriz elemental E^* y transforma E en I. Por el Teorema 6, ésta transforma a E^* en E^*E, de modo que $E^*E=I$. Como E tiene inversa a la izquierda, no es singular.

COROLARIO 2. *Si dos matrices* $m \times n$ *son equivalentes por filas, será* $B=PA$, *con* P *regular.*

Pues, por el Teorema 6, es $B=E_nE_{n-1}\ldots E_1A$, siendo las E_i elementales, y por ende, regulares.

EJERCICIOS

1. Demostrar la equivalencia por filas de $\begin{pmatrix} 5 & 2 & 7 \\ -3 & 4 & 1 \\ -1 & -2 & -3 \end{pmatrix}$ y $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

2. Reducir cada una de las siguientes matrices a la forma escalonada equivalente por filas:

a) $\begin{pmatrix} 1 & -1 & 3 \\ 2 & -4 & 1 \\ 0 & 3 & 2 \end{pmatrix}$, b) $\begin{pmatrix} -5 & 6 & -3 \\ 3 & 1 & 11 \\ 4 & -2 & 8 \end{pmatrix}$, c) $\begin{pmatrix} 1 & 6 & -2 & 5 \\ 4 & 0 & 4 & -2 \\ 7 & 2 & 0 & 2 \\ -6 & 3 & -3 & 3 \end{pmatrix}$.

d) $\begin{pmatrix} 2 & -1 & 3 & 2 \\ 0 & 2 & 1 & 4 \\ 4 & -2 & 3 & 9 \\ 2 & -3 & 4 & 5 \end{pmatrix}$, e) $\begin{pmatrix} i & 1 & -i & 1+i \\ 1 & -i & i & 2-i \\ -1 & 0 & 1 & 0 \\ 2 & i & 2i & 3i \end{pmatrix}$.

3. Hallar las matrices escalonadas equivalentes por filas de cada una de las matrices dadas en Ejerc. 9 a), Cap. VIII, § 2, y Ejerc. 4, Cap. VIII, § 3

4. a) Presentar todas las matrices elementales 3×3.
   b) Dibujar un diagrama que represente cada matriz elemental $n\times n$ de la forma (4)-(6).
5. Hallar la inversa de cada una de las matrices elementales 4×4, H_{2}, $I+2E_{3}$, F_{1}, que aparecen en el texto.
6. Demostrar el Teor. 6 por cálculo directo en el caso de las matrices 3×3.
* 7. Demostrar que cualquier operación elemental entre filas del tipo a) puede efectuarse por una sucesión de seis operaciones de los tipos b) y c). (*Sugerencia:* Probar con matrices 2×2.)

3. Equivalencia por filas y matrices inversas

Sea A una matriz $n\times n$ regular, y sea D la correspondiente matriz escalonada (Teorema 4). Por el Corolario 2 del § 2, $D=PA$ es el producto de dos matrices regulares, luego también será regular. Por lo tanto, según el Teorema 5, D debe ser triangular; todos los elementos de la diagonal principal serán iguales a 1, los elementos inferiores serán nulos y los situados encima serán cualesquiera, como aquí se indica. Estos últimos, es decir, los situados sobre la diagonal principal, pueden ser eliminados por nuevas operaciones entre las filas. Por ejemplo, los elementos de la última columna pueden ser eliminados sumando a cada fila un múltiplo conveniente de la última. El resultado final será la matriz idéntica; así hemos demostrado

$$\begin{bmatrix} 1 & d_{12} & \dots & d_{1n} \\ 0 & 1 & & d_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & 1 \end{bmatrix}$$

TEOREMA 7. *Una matriz cuadrada es regular si, y sólo si, es equivalente por filas a la matriz idéntica.*

Como las operaciones entre filas pueden llevarse a cabo con premultiplicaciones por matrices elementales E_i, cualquier matriz regular A puede reducirse a I como en

$$E_s E_{s-1} \dots E_1 A = I.$$

Multipliquemos a la derecha por A^{-1} los dos miembros de esta igualdad, y resultará

$$E_s E_{s-1} \dots E_1 I = A^{-1}. \tag{7}$$

La matriz a la izquierda es el resultado de aplicar a la identidad I la sucesión de operaciones $E_1, \dots, E_s$. Esto demuestra

TEOREMA 8. *Si una matriz cuadrada* A *se reduce a la identidad por una sucesión de operaciones entre sus filas, la misma sucesión de operaciones aplicada a la matriz identidad* I *da como resultado la matriz inversa de* A.

Resulta de aquí un método práctico para construir la inversa: dada cualquier matriz A, se puede proceder a una sucesión de operaciones racionales que darán una inversa de A o reducirán A a una matriz equivalente singular. En este último caso, A no tiene inversa. Para matrices de orden superior al 3×3, este método resulta más cómodo que el deducido de la teoría de determinantes, que es el que suele aplicarse para hallar A^{-1} (cfr. § 6).

Digamos al paso, que cualquier matriz regular P es la inversa de otra matriz regular $(P^{-1})^{-1}$; por lo tanto, puede escribirse como en (7), como un producto de matrices elementales. Combinando esto con el Corolario 1 del Teorema 6, obtenemos el siguiente resultado:

TEOREMA 9. *Una matriz cuadrada* P *es regular si, y sólo si, puede expresarse como un producto de matrices elementales.*

$$P = E_s E_{s-1} \dots E_1. \tag{8}$$

COROLARIO 1. *Dos matrices* $m \times n$, A *y* B, *son equivalentes por filas si, y sólo si,* B = PA *para alguna matriz regular* P.

Pues B es equivalente por filas a A si, y sólo si (Teor. 6), $B = E_n E_{n-1} \dots E_1 A$, siendo las E_i elementales. Y, por el Teor. 9, esto equivale a $B = PA$, con P regular.

El Teor. 9 tiene una interpretación geométrica simple en el caso de dos dimensiones. Las únicas matrices elementales 2×2 son

$$H_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad M_1 = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix},$$

$$F_{12} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad F_{21} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Las correspondientes transformaciones lineales del plano son, como en Cap. VIII, § 1,

(H_{12}) una reflexión del plano en la recta inclinada 45° que pasa por el origen;

(M_i, para c positivo) una compresión (o dilatación) paralela al eje x o al eje y;

(M_i, para c negativo) una compresión seguida de una reflexión en los ejes;

(F_{ij}) un corrimiento de cizalla sobre uno de los ejes.

Esto da

COROLARIO 2. *Cualquier transformación lineal biunívoca del plano puede representarse como producto de corrimientos sobre un eje, dilataciones (o compresiones) en una dirección y reflexiones.*

Hemos obtenido, pues, un fundamental resultado geométrico, mediante razonamientos algebraicos sobre matrices. Un resultado análogo puede lograrse para el espacio de tres o más dimensiones.

La matriz $m \times n$, A, tiene como *resultante*, según vimos en § 1, el conjunto de todas las combinaciones lineales de las filas de A. Ahora bien, cualquier operación elemental entre filas no hace más que reemplazar una fila de A por una cierta combinación lineal de todas ellas; luego el resultante engendrado por las filas de la matriz equivalente PA está contenido seguramente en el resultante de A; como la relación de equivalencia es simétrica, el resultante de A estará también contenido en el de PA; por tanto,

TEOREMA 10. *Las matrices equivalentes por filas tienen el mismo resultante y la misma característica.*

Investigación de una base. Dados m vectores en $V_n(F)$, se trata de hallar una base para el subespacio S engendrado por estos vectores. Escritos los vectores como filas de una matriz $m \times n$, A, la reduciremos a la forma escalonada D. Las filas no nulas engendran el resultante S de A, por el Teorema 10, y son independientes, por el Teorema 5.

De igual modo puede establecerse un *criterio* para establecer la *independencia* de varios vectores dados. Sean éstos m vectores de $V_n(F)$; los escribiremos como filas de una matriz $m \times n$ y la reduciremos a forma escalonada D. Los vectores serán independientes si, y sólo si, D no tiene ninguna fila de ceros. Si no son independientes, la característica de D da el número máximo de los vectores linealmente independientes en el conjunto dado.

Los métodos de reducción que acabamos de explicar, utilizando las operaciones elementales con filas, implican exclusivamente operaciones racionales en el campo de los elementos de la matriz.

Por ejemplo, si los elementos de una matriz A son números racionales, mientras el campo F es el de todos los números reales, las operaciones elementales pueden conducirse exactamente como si el campo contuviese sólo a los números racionales. En cualquier campo se llega a la misma forma escalonada y, por lo tanto, al mismo número de filas independientes.

TEOREMA 11. *Si una matriz A sobre el campo F tiene todos sus elementos pertenecientes a un campo más restringido F′, la característica de A relativa a F es la misma que la característica de A relativa al campo menos amplio F′.*

Las operaciones descritas en la equivalencia según filas son exactamente las mismas utilizadas para resolver sistemas de ecuaciones (Cap. II, § 3). Para mostrar esta analogía, consideremos m ecuaciones $\sum_j a_{ij}x_j = b_i$, con n incógnitas x_j ($i = 1, \ldots, m$; $j = 1, \ldots, n$). Los coeficientes de las incógnitas forman una matriz $m \times n$, $A = \|a_{ij}\|$, mientras que los términos constantes b_i constituyen una matriz de una columna B'. El sistema de ecuaciones puede escribirse en forma matricial $AX' = B'$, donde X' es la columna transpuesta del vector $X = (x_1, \ldots, x_n)$. Si la columna B' de constantes se agrega a la matriz A, se formará una matriz $\|A, B\|$ de tipo $m \times (n+1)$, a la cual se llama *matriz del sistema*, y es igual, como decimos, a la matriz de los coeficientes *ampliada* con la columna de los términos constantes. Las operaciones entre las filas de esta matriz ampliada se corresponden con las transformaciones del sistema dado en otro equivalente, y, por lo tanto, *dos sistemas de ecuaciones* $AX' = B'$ *y* $A^*X' = B^{*\prime}$ *tienen las mismas soluciones* X' *si sus correspondientes matrices son equivalentes por filas.*

EJERCICIOS

1. Hallar las inversas de $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 4 & 1 \\ 1 & 3 & 0 \end{pmatrix}$ y $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{pmatrix}$.
2. Hallar las inversas (si existen) de las matrices del Ejerc. 2, § 2.
3. Comprobar en V_6 la independencia de los siguientes vectores y encontrar una base para los subespacios que ellos engendran:

   a) (2, 4, 3, —1, —2, 1), (1, 1, 2, 1, 3, 1), (0, —1, 0, 3, 6, 2);

   b) (2, 1, 3, —1, 4, —1), (—1, 1, —2, 2, —3, 3), (1, 5, 0, 4, —1, 7).
4. En el Ejerc. 3, encontrar una base para el subespacio engendrado por los vectores a) y b), tomados en conjunto.

5. Resolver el Ejerc. 1 de § 4, Cap. VII, por el método de equivalencia por filas.
6. Hallar las características y bases para los resultantes de las matrices siguientes:

a) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$, b) $\begin{pmatrix} 1 & 2 & 1 & 2 \\ 3 & 2 & 3 & 2 \\ -1 & -3 & 0 & 4 \\ 0 & 4 & -1 & -3 \end{pmatrix}$, c) $\begin{pmatrix} 1 & 2 & 4 & 5 & 7 \\ 1 & 2 & 3 & 4 & 5 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}$.

7. Escribir cada una de las matrices que siguen como producto de matrices elementales:

a) $\begin{pmatrix} 3 & 6 \\ 2 & 1 \end{pmatrix}$, b) $\begin{pmatrix} 4 & -2 \\ 3 & -5 \end{pmatrix}$, c) la primera matriz de Ejerc. 1.

8. Representar la transformación $x'=2x-5y$, $y'=-3x+y$ como un producto de corrimientos, compresiones y reflexiones.
*9. Para un espacio tridimensional, establecer y demostrar algo análogo al Corol. 2 del Teor. 9. Utilizando Ejerc. 7, § 2, interpretar el resultado.
10. Demostrar que cualquier matriz 2×2 regular puede representarse como producto de las matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, y $\begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}$, donde $c \neq 0$ es un escalar cualquiera. ¿Qué significa este resultado, geométricamente?
11. Demostrar que la característica de un producto no excede nunca a la característica de cada factor.
12. Demostrar que un sistema de ecuaciones lineales $AX'=B'$ tiene una solución si, y sólo si, la característica de A es igual a la de la matriz orlada $\| A, B' \|$.
13. Sea $AX'=B'$ un sistema de ecuaciones lineales no homogéneas con una solución particular $X'=X_0'$. Demostrar que cualquier otra solución X' puede escribirse como $X'=X_0'+Y'$, siendo Y' una solución de la ecuación homogénea $AY'=0$, y recíprocamente.
14. Demostrar que si un sistema de ecuaciones lineales con coeficientes en un campo F no tiene soluciones en F, tampoco las tiene en un campo más amplio.
15. a) ¿Puede una matriz ser equivalente por filas a más de una matriz escalonada?
    b) Hallar una forma canónica de las matrices en la equivalencia por filas.

4. Equivalencia en general y formas canónicas

Las operaciones con las columnas de una matriz son análogas a las operaciones con las filas. Así, en una matriz A designaremos como *operación elemental entre columnas* cualquiera de las tres siguientes: a) el intercambio de dos columnas; b) la multiplicación de cualquier columna por un escalar no nulo, y c) la adición de una columna a otra.

Al reemplazar A por su transpuesta A', las operaciones elementales entre filas se cambian por operaciones elementales entre columnas y viceversa. En particular, A se podrá transformar en B mediante una serie de operaciones elementales entre *columnas*, cuando la transpuesta A' se pueda transformar en B' por una sucesión de operaciones elementales entre *filas*, y sólo en este caso. Aplicando al Corolario 1 del Teorema 9, esto significa que $B'=PA'$, o sea, $B=(B')'=(PA')'=AP'=AQ$, donde $Q=P'$ es regular. Inversamente, $B=AQ$ con Q regular, expresa que B es equivalente por columnas a A. Por lo tanto, la aplicación de las operaciones entre columnas es equivalente a la *post*-multiplicación por un factor no singular. Este *post*-factor puede obtenerse explícitamente aplicando las mismas operaciones a la matriz idéntica, como en el Teorema 6.

Las operaciones entre filas y entre columnas pueden ser aplicadas juntamente. Podemos decir que dos matrices $m \times n$, A y B, son equivalentes si, y sólo si, puede pasarse de A a B por una sucesión de operaciones elementales entre filas y entre columnas. Y entonces tendremos el siguiente resultado:

TEOREMA 12. *Una matriz* A *del tipo* m×n *es equivalente a otra matriz* B *si, y sólo si,* B=PAQ, *siendo* P *y* Q *matrices regulares* m×m *y* n×n *convenientemente elegidas.*

¿Hasta qué punto puede simplificarse una matriz A por operaciones entre filas y entre columnas? Si $A \neq 0$, habrá en A algún elemento no nulo b, que podemos llevar al primer lugar (vértice superior izquierdo) por conveniente intercambio de filas y columnas. Multiplicando ahora la primera fila por b^{-1} reduciremos a 1 el primer elemento. El resto de la primera columna lo convertiremos en cero restando de cada fila un múltiplo conveniente de la primera. Por el mismo método haremos que sean cero los restantes elementos de la primera fila, así que la matriz $m \times n$ tomará la forma

(9) $B=\begin{pmatrix} 1 & 0 \\ 0' & C \end{pmatrix}$, con estas dimensiones:

$$\left\{ \begin{matrix} 0 & 0' & C \\ 1\times(n-1) & (m-1)\times 1 & (m-1)\times(n-1) \end{matrix} \right\}.$$

Ahora se repetirá este proceso para el bloque C; si $C \neq 0$ se pondrá un 1 en su vértice superior izquierdo y 0 en los restantes elementos de las primeras fila y columna, y así sucesivamente.

TEOREMA 13. *Cualquier matriz* m × n, A, *es equivalente a una matriz* D *cuyos elementos no nulos son los* r *elementos* $d_{11} = d_{22} = \ldots = d_{rr} = 1$ *de la diagonal principal,*

$$(10) \quad D = \begin{pmatrix} I & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}, \quad \textit{donde} \begin{cases} I \textit{ es la identidad } r \times r, \\ 0_{ij} \textit{ es una matriz } i \times j \textit{ de ceros.} \end{cases}$$

El teorema puede demostrarse formalmente, por inducción sobre n. Para $n=1$ es trivial. Suponiendo que el teorema es cierto para matrices $(m-1) \times (n-1)$, reduciremos la matriz dada A a la forma B de (9) y aplicaremos esta hipótesis para reducir C a la forma diagonal, con lo cual A tomará la forma diagonal requerida.

¿Qué significado tiene r, número de elementos 1 en la diagonal de la matriz equivalente D? En D, las r primeras filas son, precisamente, las primeras r unidades vectoriales $I_1, \ldots, I_r$; por lo tanto, r es precisamente la característica de D (el máximo número de filas independientes). Pero la matriz original A tiene la misma característica, como vamos a probar ahora mismo.

TEOREMA 14. *Dos matrices equivalentes tienen la misma característica.*

Demostración. Sabemos ya (Teorema 10) que dos matrices equivalentes por filas tienen igual característica. Por lo tanto, sólo nos falta demostrar que dos matrices equivalentes por columnas A y $B = AQ$ (Q regular) tienen la misma característica. Además, por el Teor. 1, esto será cierto si A y B tienen la misma nulidad, lo cual se cumplirá ciertamente si ambas tienen el mismo espacio nulo. Pero $XA = O$ implica evidentemente $XB = XAQ = OQ = O$, e inversamente, $XB = O$ implica $XA = XAQQ^{-1} = XBQ^{-1} = OQ^{-1} = O$, así que *dos matrices equivalentes por columnas tienen el mismo espacio nulo* (dual del Teorema 10), lo cual completa nuestra demostración.

COROLARIO 1. *Una matriz* m × n, A, *es equivalente a una, y sólo a una, matriz diagonal de la forma (10); la característica* r *de la matriz* A *determina el número* r *de elementos 1 de la diagonal principal.*

COROLARIO 2. *Dos matrices equivalentes tienen la misma característica por columnas.*

Demostración. La característica por columnas de A (numero máximo de columnas de A linealmente independientes) es igual a la característica (por filas) de su transpuesta A'. Pero la equivalencia de A y de B impone la equivalencia de las transpuestas A' y B'. Por el teorema, A' y B' tienen la misma característica, así que A y B tendrán la misma característica por columnas.

COROLARIO 3. *La característica (por filas) de una matriz es siempre igual a su característica por columnas.*

COROLARIO 4. *Dos matrices* $m \times n$ *son equivalentes si, y sólo si, tienen ambas la misma característica.*

Si son equivalentes, tienen ambas la misma característica (Teorema 14); si tienen la misma característica, son ambas equivalentes a la misma expresión canónica D; luego lo son la una a la [illegible] Por consiguiente, la característica es un invariante de una mat en el grupo de todas las transformaciones $A \to PAQ$ (con P y Q regulares). De hecho, constituye un sistema completo de invariantes, en el sentido de Cap. IX, § 5.

EJERCICIOS

1. Comprobar el Corolario 3 del Teorema 14, calculando la característica por filas y por columnas: a) en Ejerc. 1 de § 2; b) en Ejercs. 6 a) y b) de § 3.
2. Hallar una matriz diagonal equivalente con cada matriz de Ejerc. 2, § 2.
3. Hacer lo mismo para las matrices de Ejerc. 6, § 3.
4. Sea T una transformación lineal de un espacio vectorial m-dimensional V en un n-espacio W. Mostrar que, por elección de bases convenientes en V y en W, las ecuaciones de T toman la forma $y_i = x_i$ $(i=1, \ldots, r)$, $y_j = 0$ $(j=r+1, \ldots, n)$.
5. a) Demostrar que la transpuesta de cualquier matriz elemental es elemental.
   b) Utilizando esto, dar una nueva demostración del hecho de que la transpuesta de cualquier matriz regular es regular.
*6. Si A y B son matrices $n \times n$ de características r y s, demostrar que la característica de AB nunca es menor que $(r+s)-n$. (*Sugerencia:* Utilizar la forma canónica para A.)

a) Demostrar la ley de Sylvester para la nulidad: La nulidad de un producto AB no excede a la suma de las nulidades de los factores, y nunca es menor que la nulidad de un factor si B es cuadrada.
b) Dar ejemplos que muestren que ambos límites pueden ser alcanzados por la nulidad de AB.

a) Sea A una matriz con elementos enteros y sean en ella las operaciones elementales: I) el intercambio de dos filas o columnas; II) la multiplicación de una fila (columna) por ± 1; III) la adición de una fila (columna) a otra fila (columna). Demostrar que A puede reducirse por estas operaciones a una matriz diagonal D con elementos enteros no negativos (no necesariamente todos 1) sobre la diagonal. Ilustrar este resultado para la matriz de Ejerc. 2 a), § 2.
b) Demostrar que los elementos no nulos d_{ii} de la diagonal pueden elegirse de tal modo, que cada uno sea divisor del siguiente.

Por definición, una *forma bilineal* en dos conjuntos $x_1, \ldots, x_n$ e $y_1, \ldots, y_n$ de n variables cada uno, es una suma doble $\sum_{i,j} x_i a_{ij} x_j = XAY'$. La *característica* de la forma es la característica de su matriz $A = \| a_{ij} \|$.
a) Demostrar que una forma bilineal de característica r puede reducirse por transformaciones lineales no singulares de los dos conjuntos de variables a la forma $x_1'y_1' + \ldots + x_r'y_r'$. (*Sugerencia:* Determínese primero cómo una transformación de las variables afecta a la matriz.)
b) Demostrar que dos formas bilineales son equivalentes para las transformaciones regulares de variables si, y sólo si, tienen la misma característica.

Sea $B(\xi, \eta)$ una función de dos vectores ξ, η con las propiedades $B(a_1\xi_1 + a_2\xi_2, \eta) = a_1 B(\xi_1, \eta) + a_2 B(\xi_2, \eta)$ y $B(\xi, b_1\eta_1 + b_2\eta_2) = b_1 B(\xi, \eta_1) + b_2 B(\xi, \eta_2)$. Demostrar que B puede expresarse como una forma bilineal en las coordenadas de ξ, η y demostrar que, inversamente, cualquier forma bilineal debe tener estas dos propiedades.

Demostrar que una matriz A de tipo $m \times n$ tiene una característica no mayor que 1 si, y sólo si, puede representarse como un producto $A = BC$, donde B es $m \times 1$ y C es $1 \times n$.

Demostrar que cualquier matriz de característica r es la suma de r matrices de característica 1.

Demostrar que cualquier forma bilineal de característica r es expresable como

$$\sum_i (b_{i1}x_1 + \ldots + b_{in}x_n)(c_{i1}y_1 + \ldots + c_{in}y_n), \text{ con } i = 1, \ldots, r;$$

esto es, como la suma de r productos de formas lineales.

En la teoría de equivalencia por columnas, descríbase cuidadosamente los tipos de matrices correspondientes a las matrices escalonadas en la equivalencia por columnas.

Definición del determinante y sus propiedades elementales

Generalmente no se reconoce que gran parte de la teoría de las matrices puede desarrollarse, como en los dos capítulos anteriores,

sin necesidad de emplear los determinantes. En el resto del sente capítulo estudiaremos las propiedades de las matrices cionadas con sus determinantes, utilizando también las ma elementales que hemos introducido en el anterior estudio d matrices equivalentes por columnas.

En función de los determinantes, obtendremos después la ción característica de una matriz estudiando mediante ella el nificado de la forma canónica de una transformación lineal.

La fórmula para resolver un sistema de dos ecuaciones lin lleva de modo natural a los determinantes. Dos ecuaciones $+b_1y=k_1$, $a_2x+b_2y=k_2$, tienen solución única

$$x=(k_1b_2-k_2b_1)/(a_1b_2-a_2b_1), \qquad y=(a_1k_2-a_2k_1)/(a_1b_2-a_2$$

supuesto que $a_1b_2-a_2b_1\neq 0$. Los polinomios que aparecen e merador y denominador se llaman *determinantes*,

$$(11)\qquad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2-a_2b_1, \qquad \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix} = k_1b_2-k_2b_1.$$

De modo semejante podemos resolver un sistema de tres ciones lineales $\Sigma a_{ij}x_j=k_i$. El denominador de cada solución sulta ser

$$(12)\qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}- \\ -a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}. \end{matrix}$$

En el desarrollo aparecen seis productos. En cada uno viene un factor de la primera fila, otro de la segunda y otro tercera. Cada columna está también representada en cada p to, así que un término de (12) tiene la forma $a_{1-}a_{2-}a_{3-}$, con el reemplazado por alguna permutación de los índices 1, 2, 3, q dican las columnas. De las seis permutaciones posibles, la pares, I, (123), (132), aparecen en los productos precedidos y las tres impares en los productos asociados con el signo n La experiencia mostraría inmediatamente que el mismo pri se puede aplicar a la solución del sistema de n ecuaciones incógnitas.

El determinante de una matriz $n\times n$, $A=\|a_{ij}\|$, será una de monomios $\pm a_{1-}a_{2-}\dots a_{n-}$, donde los guiones se reemplazan guna permutación ϕ de los subíndices. Si esta ϕ transforma i

el término correspondiente puede escribirse $\pm a(1, 1\phi) \ldots a(n,$ poniendo $a(i, j)$ en vez de a_{ij}; en estos términos interviene e mente un factor de cada fila y uno de cada columna. El sig que se elige es el llamado *sg* ϕ (esto es: *signo* de ϕ).

DEFINICIÓN. *Determinante* $|A|$ *de una matriz* $A = \|a_{ij}$ *tipo* $n \times n$, *es el siguiente polinomio de los elementos* (*) $a_{ij} = a$

$$(13) \qquad det\ (A) = |A| = \sum_{\phi} (sg\ \phi) a(1, 1\phi) a(2, 2\phi) \ldots a(n, n\phi).$$

Este polinomio es una suma de $n!$ *términos, uno para cada mutación* $i \to i\phi$ *entre los números* $1, 2, \ldots, n$. *El término qu rresponde a la permutación* ϕ *es un producto de* n *factores, u cada fila de* A; *el factor* $a(i, i\phi)$ *pertenece a la fila* i *y a la c na* $i\phi$. *El signo que precede a cada término es el de* $sg\,\phi$, *es un factor* $+1$ *o* -1, *según* ϕ *sea una permutación par o impar pítulo VI, § 10).*

Se llama orden de un determinante, al orden n *de su matr*

Cada columna interviene una y sólo una vez en cada tér de $|A|$, lo cual significa que $|A|$ es función lineal y homogén los elementos $a_{i1}, \ldots, a_{in}$ de la fila i-ésima de A. Agrupand términos en que interviene cada a_{ij} obtendremos una expresió

$$(14) \qquad |A| = A_{i1}a_{i1} + A_{i2}a_{i2} + \ldots + A_{in}a_{in},$$

en la que el coeficiente A_{ij} de a_{ij} es llamado el *adjunto* (o *cofa* de a_{ij}; su expresión será, pues, un polinomio con los element las restantes columnas de A. Este adjunto puede también defi como la derivada parcial $A_{ij} = \partial |A| / \partial a_{ij}$. Como en cada factor interviene cada fila y cada columna una sola vez, es claro qu el adjunto A_{ij} no pueden intervenir ni los elementos de la fila los de la columna j. En él intervienen solamente los elemento «menor» o submatriz M_{ij}, que es la matriz obtenida al sup de A la fila i y la columna j.

Las filas y columnas intervienen simétricamente en $|A|$:

TEOREMA 15. *Si* A' *es la transpuesta de* A, $|A'| = |A|$.

(*) Los elementos $a(i, j)$ pertenecen a un campo F o, más generalmente, anillo conmutativo con unidad.

Demostración. El elemento $a_{ij}'=a_{ji}$ de A' se obtiene invirtiendo los subíndices. Un término de $|A|$, con $j=i\phi$ e $i=j\phi^{-1}$ es

$$(sg\,\phi)\prod_i a(i,i\phi)=(sg\,\phi)\prod_j a(j\phi^{-1},j)=(sg\,\phi)\prod_j a'(j,j\phi^{-1}).$$

Resulta, pues, un término del polinomio $|A'|$, cuya correspondiente permutación es la inversa ϕ^{-1} de la permutación ϕ. Luego es $(sg\,\phi)=(sg\,\phi^{-1})$, ya que ϕ es par si, y sólo si, su inversa ϕ^{-1} es también par (Cap. VI, § 10). Por lo tanto, $A=A'$, c. q. d.

¿Qué efecto producirán sobre un determinante las operaciones elementales entre filas?

Regla 1. Si la columna i-ésima de A se multiplica por un escalar $c\neq 0$, el determinante $|A|$ queda multiplicado por c. Pues, en la expresión lineal homogénea (14), un factor extra c en cada término $a_{i1}, \ldots, a_{in}$ equivale, simplemente, al mismo factor extra en $|A|$.

Regla 2. Al permutar dos filas de A, cambia el signo de $|A|$. Por la simetría (Teorema 14), bastará probar que el intercambio de dos columnas cambia el signo. Este intercambio estará representado por una permutación impar ϕ_0 en los índices de las columnas; así se reemplazará A por $B=\|b_{ij}\|$, donde $b(i,j)=a(i,j\phi_0)$. Entonces,

$$|B|=\sum_\phi (sg\,\phi)\prod_i b(i,i\phi)=\sum_\phi (sg\,\phi)\prod_i a(i;i\phi\phi_0).$$

Como las permutaciones forman grupo, los productos $\phi\phi_0$ (con ϕ_0 fijo) incluyen todas las permutaciones, así que en $|B|$ figuran también todos los términos de $|A|$. Pero los signos de los términos están cambiados, pues siendo ϕ_0 impar, $\phi\phi_0$ será par si ϕ es impar, y viceversa ($sg\,\phi\phi_0=-sg\,\phi$). Esto demuestra la Regla 2.

LEMA. *Si* A *tiene dos filas iguales*, $|A|=0$.

Demostración. El intercambio de estas dos filas no altera a $|A|$, por ser ambas iguales, pero cambia el signo de $|A|$, por la Regla 2. Luego $|A|=-|A|$, $|A|+|A|=0$ y $|A|=0$. En un campo en que $1+1=0$, esta demostración falla, pero el lema es igualmente cierto (ver el Apéndice a esta sección).

Para las consideraciones sobre los adjuntos (§ 6), es conveniente expresar este lema mediante una igualdad. En A, reemplacemos la fila i por la fila k. En tal caso, dos filas resultan iguales y el deter-

minante será nulo. Pero este determinante puede expresarse cambiando la fila i por la fila k en la expresión lineal homogénea (14), así que

$$0 = A_{i1}a_{k1} + A_{i2}a_{k2} + \ldots + A_{in}a_{kn} \qquad (i \neq k). \tag{15}$$

Regla 3. Si a la fila i se le agrega c veces la fila k, $|A|$ no varía. Esta operación reemplaza cada a_{ij} por $a_{ij} + ca_{kj}$; por la expresión lineal homogénea (14), el nuevo determinante es

$$\sum_j A_{ij}(a_{ij} + ca_{kj}) = \sum_j A_{ij}a_{ij} + c\sum_j A_{ij}a_{kj} = |A| + 0,$$

por (14) y (15); luego el determinante $|A|$ queda inalterado.

Estas reglas pueden sumariarse mediante las matrices elementales. Una operación elemental transforma la identidad I en una matriz elemental E, y A en el producto EA. El determinante $|I| = 1$ es a su vez cambiado en $|E| = c$, -1 o 1 (Reglas 1, 2 o 3), mientras que $|A|$ se transforma en $|EA| = c|A|$, $(-1)|A|$ o $|A|$, respectivamente. Esto demuestra que $|EA| = |E||A|$; por simetría (Teorema 15), lo mismo se aplica al postfactor E.

TEOREMA 16. *Si* E *es una matriz elemental,*

$$|EA| = |E||A| = |AE|.$$

Estas reglas dan un método para calcular un determinante $|A|$. Se reduce A por operaciones elementales a la forma diagonal D; llamemos t al número de intercambios de filas que se hayan verificado y $c_1, \ldots, c_s$ a los distintos escalares por los que se han multiplicado las filas (o columnas) de A. La matriz diagonal D (de ceros y unos) tiene $|D| = 0$ o 1; por el Teorema 16, será, pues, $|A| = (-1)^t(c_1 \ldots c_s)^{-1}|D|$.

Otro método de cálculo es acudir al desarrollo (14) mediante las submatrices M_{ij} antes mencionadas.

Regla 4. $A_{ij} = (-1)^{i+j}|M_{ij}|$; en palabras: cada adjunto A_{ij} es igual al determinante de la correspondiente submatriz M_{ij} precedido del signo de $(-1)^{i+j}$. Este signo se halla en la posición (i, j) distribuyendo los $(\pm)$ como las blancas y negras de un tablero de ajedrez, estando la casilla superior izquierda ocupada por un $+$. Demostremos primero esta regla para $i = j = 1$. La definición (13) muestra claramente que los términos en que interviene a_{11} son

exactamente los términos que resultan de todas las permutaciones φ en las que $1\varphi=1$. Una permutación par (impar) de este tipo continúa siendo una permutación par (impar) de los restantes índices $2, \ldots, n$, así que es factor común de los términos que constituyen, exactamente, el desarrollo de $|M_{11}|$. Cualquier otro adjunto A_{ij} puede ahora reducirse a este caso especial, llevando el término a_{ij} a la posición (1, 1) por medio de $i-1$ cambios con las filas superiores y $j-1$ cambios con las columnas a la izquierda. Estas operaciones no alteran $|M_{ij}|$, pues la posición relativa de las filas y columnas de M_{ij} queda inalterada, pero cambia $i+j-1-1$ veces el signo de $|A|$ y por lo tanto el signo del adjunto de a_{ij}. Esto demuestra la regla.

Un caso de especial interés es aquel en el cual todos los elementos de la primera fila, a excepción del primero, son nulos. En el desarrollo (14) aparece entonces solamente el primer adjunto $|M_{11}|=A_{11}$, como sigue:

$$(16) \qquad \begin{vmatrix} c & 0 \\ K & B \end{vmatrix} = c\,|B| \qquad \begin{Bmatrix} 0 & K & B \\ 1\times n & (n-1)\times 1 & (n-1)\times(n-1) \end{Bmatrix}.$$

Por esta regla se demuestra fácilmente que *el determinante de una matriz diagonal es el producto de sus elementos diagonales*. (Lo cual es una consecuencia directa de la definición.)

Apéndice. También para un campo con $1+1=0$ es válido el lema precedente, como mostraremos para el caso especial del campo J_2 de los enteros módulo 2. Sea A una mátriz de indeterminadas x_{ij} independientes, con dos filas iguales. Entonces $|A|$ puede escribirse como un polinomio con coeficientes *enteros* $(1+1\neq 0)$; por el lema, $|A|=0$. *Ahora* reduzcamos los coeficientes al módulo 2. Como $|A|=0$ idénticamente, todos los coeficientes eran cero antes de esta reducción, luego también serán cero después de la reducción, c. q. d. Aquí aparece la potencia del concepto de homomorfismo. Un razonamiento semejante se hará siempre que $1+1=0$.

EJERCICIOS

1. Si una matriz $n\times n$ tiene más de n^2-n elementos iguales a 0, su determinante es nulo.
2. Calcular los determinantes de las matrices de Ejerc. 2, § 2.

3. a) Si $A=\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$, calcular $|A|$ por los menores de la primera fila y por los menores de la primera columna, y comparar los resultados.
   b) Calcular $|A|$ en el supuesto de que los elementos de A son enteros mód. 2.
4. Escribir los términos positivos en el desarrollo de un determinante general 4×4.
5. Si n es impar y $1+1\neq 0$, mostrar que cualquier matriz hemisimétrica $n\times n$ tiene determinante 0.
6. a) Deducir el siguiente desarrollo del «determinante de Vandermonde»:

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_1-x_2)(x_2-x_3)(x_3-x_1);$$

   b) Generalizar este resultado al caso $n\times n$.
7. Hallar el desarrollo del determinante del Ejerc. 6: a) si cada x_i^2 en la tercera columna se reemplaza por x_i^3; b) generalizar.
8. Demostrar que el determinante de cualquier matriz de permutación es ± 1.
9. Demostrar que el determinante de una matriz monomial es el producto de los elementos no nulos por ± 1.
10. a) En el plano, mostrar que la línea que une el punto (a_1, a_2) al punto (b_1, b_2) tiene la ecuación

$$\begin{vmatrix} x_1 & x_2 & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{vmatrix} = 0.$$

   b) Generalizar este resultado a espacios de más dimensiones.
*11. a) Si todo elemento a_{ij} de la matriz A es una función de x, demostrar que

$$\frac{d\,|A|}{dx} = \sum_{j,k=1}^{n} \frac{da_{jk}}{dx} A_{jk}.$$

   b) Utilizar esto para demostrar que $A_{ij} = \dfrac{\partial |A|}{\partial a_{ij}}$.
*12. Si A y C son matrices cuadradas, demostrar que $\begin{vmatrix} A & B \\ 0 & C \end{vmatrix} = |A|\cdot|C|$.

6. Producto de determinantes

Mediante las operaciones elementales entre filas y columnas, cualquier matriz cuadrada A resulta equivalente a una matriz diagonal D (Teorema 13), así como A puede obtenerse a partir de D, por premultiplicación y postmultiplicación por matrices elementales E_i y $E^{(i)}$, como en (7),

$$A = E_s \ldots E_1 D E^{(1)} \ldots E^{(t)}. \tag{17}$$

Las reglas $|EA|=|E|\cdot|A|$ y $|AE|=|A|\cdot|E|$ del Teorema 16 muestran que el determinante del producto (17) resulta tomando sin más el determinante de los factores; de modo que se obtendrá

$$(18) \qquad |A|=|E_s|\ldots|E_1||D||E^{(1)}|\ldots|E^{(t)}|.$$

Como cada $|E_i|\neq 0$, será el determinante $|A|\neq 0$ si, y sólo si, $|D|\neq 0$. La forma canónica D tiene exactamente r elementos 1 en la diagonal principal, siendo r la característica de A, mientras que el determinante $|D|$ es el producto de sus n elementos diagonales. Por lo tanto, $|D|\neq 0$ si, y sólo si, $r=n$; esto es, si, y sólo si, A es regular. Por lo tanto, (18) demuestra:

TEOREMA 17. *Una matriz cuadrada* A *es regular si, y sólo si,* $|A|\neq 0$.

Una matriz regular A es un producto $A=E_s\ldots E_1$ de matrices elementales. Si $B=E_s^*\ldots E_1^*$ es otra matriz regular, al producto AB corresponde un determinante que puede ser calculado como en (18), resultando

$$|AB|=|E_s\ldots E_1E_s^*\ldots E_1^*|=|E_s|\ldots|E_1|\cdot|E_s^*|\ldots|E_1^*|=|A|\cdot|B|.$$

TEOREMA 18. *El determinante de una matriz producto es el producto de los determinantes:* $|AB|=|A|\cdot|B|$.

El cálculo anterior demuestra esto cuando A y B son regulares. Pero si A (o B) es singular, también lo es AB, y los dos miembros de la igualdad $|AB|=|A|\cdot|B|$ son cero, c. q. d.

La inversa de una matriz A con determinante $|A|\neq 0$ existe, y puede construirse explícitamente utilizando los adjuntos en A. Las ecuaciones (14) y (15), en que intervienen estos adjuntos, pueden escribirse así:

$$(19) \qquad a_{k1}A_{i1}+\ldots+a_{kn}A_{in}=\delta_{ki}|A|, \text{ donde } \delta_{ki}=\begin{cases}1 & \text{si } i=k\\ 0 & \text{si } i\neq k\end{cases}.$$

El número δ_{ki} es precisamente el elemento (k, i) de la matriz idéntica $I=\|\delta_{ki}\|$. La igualdad (19) es muy semejante a un producto matricial; si los subíndices de los adjuntos A_{ij} se intercambian, el primer miembro de (19) da el elemento (k, i) del producto de $A=\|a_{kj}\|$ por la matriz transpuesta de adjuntos. El segundo

miembro de (19) es el elemento (k, i) de la identidad multiplicada por un escalar $|A|$; es decir,

$$(20) \qquad A \| A_{ij} \|' = |A| I.$$

La matriz $\| A_{ij} \|'$ que aparece en esta igualdad es la matriz *transpuesta* de los adjuntos de los elementos de A y se le llama matriz *adjunta* de A. Si $|A|=1$, la (20) demuestra que la adjunta es la inversa de A; en general, si $|A| \neq 0$, la (20) demuestra el

TEOREMA 19. *Si* $|A| \neq 0$, *la inversa es* $A^{-1} = |A|^{-1} \| A_{ij} \|'$.

Esta fórmula es especialmente manejable para las matrices 2×2.

La regla de Cramer, para resolver un sistema de n ecuaciones lineales con n incógnitas, es una consecuencia de esta expresión de la inversa. Tal sistema tiene la forma general $\sum_j a_{ij} x_j = b_i$, donde i y j toman valores de 1 a n. En notación matricial, el sistema se escribirá: $AX' = B'$, donde B' significa la matriz de una columna $(b_1, \ldots, b_n)'$. Si $|A| \neq 0$, la ecuación $AX' = B'$ premultiplicada por A^{-1} da la única matriz solución $X' = A^{-1}B'$. Para formular esta solución, observaremos que el elemento (i, j) en la inversa A^{-1} es precisamente $A_{ji}/|A|$. Por lo tanto,

TEOREMA 20 (Regla de Cramer). *Si el sistema de* n *ecuaciones con* n *incógnitas* $\sum_j a_{ij} x_j = b_i$ *tiene regular la matriz* $A = \| a_{ij} \|$ *de sus coeficientes, hay una solución única dada por*

$$(21) \qquad x_j = (A_{1j} b_1 + \ldots + A_{nj} b_n)/|A|, \qquad j = 1, \ldots, n,$$

donde A_{ij} *es el adjunto en* A *del elemento* a_{ij}.

El numerador de esta fórmula puede escribirse también en forma de determinante, pues es el desarrollo del determinante de la matriz que resulta sustituyendo la columna j de A por la columna de las constantes b_i. Obsérvese además que los sistemas con muchas ecuaciones se resuelven más fácilmente reduciendo la matriz a la forma «escalonada» equivalente, como en § 3.

La regla de Cramer se aplica a cualquier campo; en particular, a las ecuaciones discutidas en Cap. II, § 3 (cfr. Ejerc. 9).

Apéndice. Determinantes y característica. Se llama submatriz (o menor) de una matriz rectangular A, cualquier matriz obte-

nida a partir de A suprimiendo ciertas filas y columnas (en lo cual se incluye la posibilidad de que ninguna fila o ninguna columna sean suprimidas). La «característica por determinantes» de una matriz A puede definirse como el número máximo d de columnas que intervienen en los menores de A cuyo determinante no es nulo; dicho de otro modo, d tiene estas propiedades: 1) A tiene por lo menos un menor $d \times d$, tal como M, con $|M| \neq 0$; 2) Si $h > d$, cualquier menor $h \times h$ de A, tal como N, tiene $|N| = 0$.

TEOREMA 21. *La característica* r *de una matriz es igual a su característica por determinantes.*

Demostración. Supongamos primero que M es un menor $d \times d$ con $|M| \neq 0$. Como la matriz M es regular (Teorema 17), sus filas son linealmente independientes (Teor. 7, Cor. 3, Cap. VIII). Las correspondientes (y posiblemente más largas) filas de la matriz A deberán ser también linealmente independientes, pues cualquier relación lineal entre las filas de A implicaría ciertamente la relación análoga entre las filas de M. Como se han encontrado d filas independientes en A, la característica r (número máximo de filas independientes) satisface a $r \geqslant d$.

Inversamente, elijamos r filas independientes en A; éstas forman un menor $r \times n$ M_1 de la matriz $m \times n$ A. Como la característica por filas es igual a la característica por columnas (número de columnas independientes), deberá haber r columnas independientes en M_1. Estas columnas forman un menor $r \times r$ de M_1, que llamamos M_2. Sus columnas son independientes, luego M_2 es regular (Teor. VII del Cap VIII) y tiene un determinante $|M_2| \neq 0$. La construcción de este menor muestra que la característica por determinantes d es al menos igual a r. Combinando estas dos desigualdades resulta la igualdad deseada, d=r.

De estos resultados se puede obtener un *criterio para establecer la independencia (lineal) de varios vectores dados*. Sean m vectores dados por sus coordenadas $(a_{i1}, \ldots, a_{in})$ en el espacio $V_n(F)$, los cuales se pueden considerar como las filas de una matriz A de tipo $m \times n$. Si $m > n$, los vectores son ciertamente dependientes. Si $m \leqslant n$, calcularemos los determinantes de las submatrices $m \times m$ de A. Si *alguno* de estos determinantes no es nulo, la característica (número de filas independientes en A) es m; luego los vectores dados son independientes. Si todos los determinantes $m \times m$ son

nulos, la característica es menor que m. En este caso, la característica (dimensión del subespacio engendrado por los vectores dados) puede hallarse explícitamente por el cálculo sistemático de ulteriores subdeterminantes.

EJERCICIOS

1. Escribir la adjunta de la matriz 2×2 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y los productos de A por su adjunta.
2. a) Calcular la adjunta de la matriz del Ejerc. 2 a), § 2, y verificar en cada caso la regla del producto de una matriz por su adjunta.
   b) Lo mismo para la matriz del Ejerc. 2 b), § 2.
3. Por el método de las adjuntas, hallar las inversas de las matrices 4×4 elementales H_{24}, $I+2E_{31}$ y F_{12} de § 2.
4. Hallar las inversas de Ejerc. 1, § 3, por el método antedicho.
5. Si A es regular, demostrar que $|A^{-1}|=|A|^{-1}$.
6. Demostrar que el producto de una matriz singular por su adjunta es la matriz 0.
7. Demostrar que los determinantes pueden multiplicarse «columnas por columnas».
8. Escribir la regla de Cramer para tres ecuaciones con tres incógnitas.
9. Resolver las congruencias simultáneas de Ejerc. 1, § 3, Cap. II, por la regla de Cramer.
10. a) Mostrar que el par de ecuaciones homogéneas y lineales

$$a_1x+b_1y+c_1z=0, \quad a_2x+b_2y+c_2z=0$$

tiene la solución

$$x=\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad y=\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}, \quad z=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

   b) ¿Cuándo es tal solución una base del conjunto de todas las soluciones?
   c) Deducir fórmulas similares para tres ecuaciones con 4 incógnitas.
11. Demostrar que el determinante de una matriz ortogonal es ± 1.

Ejercicios sobre el Apéndice

12. Calcular la característica del determinante de las matrices de los Ejercicios 2 a) y b), § 2.
13. Comprobar por determinantes la independencia de los vectores de Ejercicio 3, § 3.
14. Mostrar directamente, partiendo de la definición de característica de un determinante, que una operación elemental entre columnas no altera la característica.

*15. a) Si A y B son matrices 3×3, mostrar que el determinante de cualquier submatriz 2×2 de AB es la suma de varios términos, cada uno de los cuales es un producto del determinante de una submatriz 2×2 de A por el de una submatriz 2×2 de B.
   b) Generalizar estos resultados y utilizarlos para demostrar que característica $(AB) \leq$ característica A.

*16. Si una matriz $n\times n$ A tiene característica r, demostrar que la característica s de la adjunta de A se determina como sigue: Si $r=n$, es $s=n$; si $r=n-1$, es $s=1$; si $r<n-1$, es $s=0$.
*17. Mostrar que el determinante de la matriz adjunta de A es $|A|^{n-1}$.
*18. Demostrar que el adjunto del adjunto de A es $|A|^{n-2}A$.

* 7. El determinante como medida de un volumen.

Los determinantes de las matrices reales $n\times n$ pueden interpretarse como volúmenes en un espacio euclídeo n-dimensional. Esta conexión viene sugerida por la fórmula para el área de un paralelogramo.

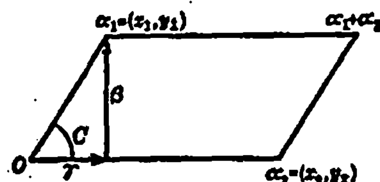


Figura 1

Cada matriz 2×2 real A, con filas α_1 y α_2, puede ser representada por un paralelogramo de vértices

$$0,\ \alpha_1,\ \alpha_2,\ \alpha_1+\alpha_2;$$

e inversamente, todo paralelogramo análogo determina una matriz (cfr. fig. 1). El área del paralelogramo es

$$\text{(22)}\qquad \text{base}\times\text{altura}=|\alpha_1|\cdot|\alpha_2|\cdot|\,\text{sen}\ C\,|,$$

donde C designa el ángulo entre los vectores α_1 y α_2 dados. Por la fórmula del coseno, (28) de Cap. VII, el *cuadrado* del área es

$$(\alpha_1,\alpha_1)(\alpha_2,\alpha_2)(1-\cos^2 C)=(\alpha_1,\alpha_1)(\alpha_2,\alpha_2)-(\alpha_1,\alpha_2)(\alpha_2,\alpha_1).$$

Este resultado aparece muy semejante al determinante de una matriz 2×2. Es, en efecto, el determinante de $\|(\alpha_i,\alpha_j)\|=AA'$.

Una fórmula semejante vale para los paralelogramos en un espacio euclídeo de cualquier número de dimensiones, y puede también extenderse a figuras m-dimensionales análogas a los paralelogramos en el espacio euclídeo de n dimensiones. Tales son los llamados *paralelepípedos*.

Para establecer esta generalización, sea A una matriz $m\times n$, con filas $\alpha_1, \ldots, \alpha_m$. Estas filas representan vectores partiendo del origen en el espacio euclídeo n-dimensional E_n. El paralelepípedo Π de E_n engendrado por los m vectores α_i consiste en todos los vectores (esto es, en sus extremos) de la forma

$$t_1\alpha_1+\ldots+t_m\alpha_m,\qquad (0\leqslant t_i\leqslant 1,\quad i=1,\ \ldots,\ m).$$

(¡Representar esto en el caso $m=n=3$; el resultado es equivalente afín a un cubo!) Esta construcción establece una correspondencia entre las matrices reales $m\times n$ y los paralelepípedos m-dimensionales en E_n. Las α_i son llamadas *aristas* del paralelepípedo Π.

El *volumen* m-dimensional de esta figura, que indicamos por $V(\Pi)$, puede definirse por inducción sobre m (incluyendo las longitudes, si $m=1$, y las áreas, si $m=2$). Al paralelepípedo de aristas $\alpha_2, \ldots, \alpha_m$ le llamaremos *base* de Π. La *altura* será la componente de α_1 ortogonal a $\alpha_2, \ldots, \alpha_m$; ella resulta expresando α_1 como la suma de una componente γ en el espacio S_{m-1} engendrado por $\alpha_2, \ldots, \alpha_m$ y una componente β ortogonal a S_{m-1} (ver fig. 1; la posibilidad de hacer esto en el caso general viene de Cap. VII, § 9):

$$(23) \qquad \alpha_1=\beta+\gamma \qquad \beta\perp S_{m-1}, \qquad \gamma \text{ en } S_{m-1}.$$

El volumen de Π se define como el *producto* del volumen $(m-1)$-dimensional de la *base* por la longitud β de la *altura*. (Veremos en seguida que el volumen así definido no depende de cuáles son las $m-1$ aristas elegidas para constituir la base.)

TEOREMA 22. *El cuadrado del volumen del paralelepípedo de aristas $\alpha_1, \ldots, \alpha_m$ es el determinante $|AA'|$, donde A es la matriz en que las coordenadas de α_i constituyen la fila i-ésima* (*).

Demostración. Como A es una matriz $m\times n$, el producto AA' es una matriz cuadrada $m\times m$. Razonemos ahora por inducción sobre m. Si $m=1$, la matriz A es una fila, y el «producto interno» $AA'=(\alpha_1, \alpha_1)$ es el cuadrado de la longitud, como afirmábamos. Supongamos el teorema cierto para matrices de $(m-1)$ filas, y consideremos el caso de m filas. Como en (23), la primera fila A_1 puede escribirse $A_1=B_1+C_1$, donde la «altura» B_1 es ortogonal a cada fila $A_2, \ldots, A_m$ ($B_1A_i'=0$), mientras que $C=c_2A_2+\ldots+c_mA_m$ es combinación lineal de ellas. Restemos sucesivamente c_i veces la fila i de la primera fila de A. Así se cambiará A en una nueva matriz A^* cuya primera fila es B_1; ahora bien, las operaciones elementales con filas suponen premultiplicar A por matrices elementales de determinantes 1, luego $A^*=PA$, con $|P|=1$ y $|A^*A^{*\prime}|=$

(*) En todo el § 7 se supone que las coordenadas de cualquier vector se toman en relación a una base ortogonal normal fija.

$=|PAA^TP^T|=|P||AA^T||P^T|=|AA^T|$. Pero si D es el bloque compuesto por las $m-1$ filas $A_2, \ldots, A_m$ de A^*, será

$$A^*A^{*T}=\begin{pmatrix} B_1 \\ D \end{pmatrix}(B_1^T, D^T)=\begin{pmatrix} B_1B_1^T & B_1D^T \\ DB_1^T & DD^T \end{pmatrix}=\begin{pmatrix} B_1B_1^T & 0 \\ 0 & DD^T \end{pmatrix}$$

donde $B_1D^T=0$ porque $B_1A_i^T=0$ para cada fila A_i de D. Por (16), el determinante es

$$|AA^T|=|A^*A^{*T}|=(B_1B_1^T)|DD^T|.$$

Como D es la matriz cuyas filas $A_2, \ldots, A_m$ engendran la base de H, $|DD^T|$ expresará el cuadrado del volumen de la base, por la hipótesis de la inducción. Además, el escalar $B_1B_1^T$ es el cuadrado de la longitud de la altura, y así tenemos la deseada expresión, mediante AA^T, de la fórmula *base × altura*.

En el caso particular de que el número de filas sea n, $|AA^T|=|A||A^T|=|A|^2$, y así demostramos (*).

Teorema 23. *Sea A una matriz real n×n de filas $\alpha_1, \ldots, \alpha_n$. El determinante $|A|$ es (excepto tal vez el signo) igual al volumen del paralelepípedo de E_n que tiene como aristas los vectores $\alpha_1, \ldots, \alpha_n$.*

El valor absoluto de un determinante no se altera por permutar las filas entre sí, de modo que el teorema demuestra que nuestra definición de volumen es independiente del orden de las aristas.

Teorema 24. *Una transformación lineal $Y=XP$ de un espacio vectorial n-dimensional euclídeo, multiplica el volumen de cualquier paralelepípedo n-dimensional por el factor $\pm|P|$.*

Demostración. Consideremos un paralelepípedo cuyas n aristas tengan respectivamente coordenadas $A_1, \ldots, A_n$. Las filas $A_1, \ldots, A_n$ se transformarán en $A_1P, \ldots, A_nP$; la matriz con estas nuevas filas será, simplemente, la matriz producto AP, siendo A la matriz de filas $A_1, \ldots, A_n$. El nuevo volumen será, pues, $|AP|=|A||P|$, siendo $|A|$ el volumen primitivo.

De aquí se deduce que la transformación $Y=XP$ conserva los volúmenes y su signo si, y sólo si, su matriz P tiene $|P|=\pm1$. El

(*) La línea directriz de esta demostración nos fué sugerida por el profesor J. S. Frame.

conjunto de todas las matrices, o de todas las transformaciones con esta propiedad, constituye el llamado *grupo unimodular*. A veces se amplía el grupo, para incluir en él a todas las P con $|P|=\pm 1$ (con lo que se constituye el grupo de todas las transformaciones que conservan el valor absoluto del volumen).

Por otra parte, el volumen de cualquier figura f en el espacio E_n puede definirse a partir del volumen del paralelepípedo, como sigue: Llamaremos *cobertura* de f a todo conjunto de paralelepípedos $\Pi_1, \ldots, \Pi_s$ que incluyen entre sus puntos a todos los puntos de f. El volumen de f es entonces el *extremo inferior* (Cap. III) de las sumas $\sum V(\Pi_i)$ de las diferentes coberturas de f. Hablando con más sencillez, es el mínimo de los volúmenes de todas las figuras circunscritas que pueden descomponerse en paralelepípedos. (Este método es el empleado usualmente en cálculo integral, al descomponer una figura en ortoedros de caras paralelas a los ejes de coordenadas.)

Ahora bien, una transformación lineal P multiplica todas las sumas $\sum V(\Pi_i)$ por el factor $|P|$, según el Teorema 24. Por lo tanto, también su límite inferior queda multiplicado por el mismo factor. Así concluimos que P transforma todos los volúmenes multiplicándolos por $|P|$. Se prueba también fácilmente que los volúmenes no se alteran por una traslación. Entonces, el Teorema 24 significa que una transformación afín $Y=XP+K$ (transformación lineal seguida de una traslación) multiplica los volúmenes por un factor constante $|P|$.

Este resultado tiene diversas aplicaciones elementales. Por ejemplo, la simetría demuestra que las tres medianas de un triángulo equilátero lo dividen en seis partes de igual área. Cualquier triángulo es equivalente en el grupo afín a uno equilátero (Cap. IX, § 14); por lo tanto, las medianas de cualquier triángulo lo dividen en seis partes de igual área.

EJERCICIOS

1. a) Calcular el área del paralelogramo con vértices (0, 0), (3, 1), (1, 3) y (4, 4) en el plano.
   b) Lo mismo para el paralelepípedo del espacio en que (0, 2, 0), (2, 0, 0), (1, 1, 5) y (0, 0, 0) son vértices adyacentes.
2. a) Si A es una matriz real $m \times n$, demostrar, como en Teorema 22, que $|AA^T| \geq 0$. Mostrar que en el caso $m=2$, este resultado es la desigualdad de Schwarz de Cap. VII, Teor. 10.

b) Mostrar que el área de un triángulo con vértices $(0, 0, 0)$, (x_1, y_1, z_1) y (x_2, y_2, z_2) es $(1/2)|AA^T|^{1/2}$ con $A = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$.

*c) El volumen de un tetraedro con tres aristas unidad sobre los ejes x, y y z, es $1/6$. Demostrar que el volumen de un tetraedro con vértices $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, es $(1/6)|BB^T|^{1/2}$, siendo B la matriz 3×3 de filas $(\alpha_2 - \alpha_1)$, $(\alpha_3 - \alpha_1)$, $(\alpha_4 - \alpha_1)$.

*d) Generalizarlo a n dimensiones.

3. Demostrar que las diagonales de cualquier paralelogramo le dividen en cuatro partes de igual área.

4. a) Si P es la intersección de las diagonales de un paralelogramo, demostrar que cualquier línea que pase por P biseca el área del paralelogramo.

b) Extender este resultado a tres dimensiones.

5. Encontrar tres planos que dividan a un tetraedro en seis partes de igual volumen.

6. En el caso de un determinante 3×3, interpretar geométricamente las operaciones elementales utilizadas en la demostración de Teorema 27.

7. a) Si m vectores $\alpha_1, \ldots, \alpha_m$ en E_n son linealmente dependientes, demostrar que el paralelepípedo que engendran tiene volumen m-dimensional igual a cero.

b) Enunciar y demostrar el recíproco.

*8. En el grupo de las matrices ortogonales, mostrar que las matrices con $|A| = +1$ (matrices «propiamente ortogonales») forman un subgrupo normal de índice 2.

9. a) Mostrar que la correspondencia $A \to |A|$ representa homomórficamente al grupo lineal sobre el grupo multiplicativo de los escalares no nulos.

b) Deducir de aquí que el grupo unimodular es un subgrupo normal del grupo lineal.

c) ¿Es un subgrupo normal del grupo lineal el grupo unimodular ampliado (esto es, el de las P con $|P| = \pm 1$)?

10. a) Demostrar que si A es una matriz con filas α_i, entonces AA^T es la matriz del producto interno $\|(\alpha_i, \alpha_j)\|$.

b) Utilizando a), demostrar que, si las α_i son ortogonales, se tiene

$$|AA^T| = |\alpha_1|^2 \cdots |\alpha_m|^2.$$

*11. Hallar el volumen común a los dos cilindros $(x+y)^2 + z^2 = a^2$; $y^2 + z^2 = a^2$, que se penetran. (*Sugerencia*. Simplificar la figura antes de integrar.)

*12. Sea $-K \leq a_{ij} \leq K$ para $i, j = 1, \ldots, n$.

a) Demostrar que si $\alpha_i = (a_{i1}, \ldots, a_{in})$, será $|\alpha_i| \leq K\sqrt{n}$.

b) Deducir $|A| \leq |\alpha_1| \cdot |\alpha_2| \cdots |\alpha_n| \leq K^n n^{n/2}$ (Teorema de Hadamard).

8. Matrices semejantes

Una misma transformación lineal T puede ser representada por diferentes matrices, según la elección de las coordenadas. Así

el plano, la transformación $(x, y) \to (2x+y, x+2y)$ es usualmente representada por la matriz A situada a la izquierda:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Pero respecto a las coordenadas $x^* = x+y$, $y^* = x-y$, la transformación resulta ser $(x^*, y^*) \to (3x^*, y^*)$, representada por la sencilla matriz diagonal D que está a la derecha.

Para expresar esta relación entre las matrices A y D se dirá que son semejantes, de acuerdo con lo que sigue:

DEFINICIÓN. *Dos matrices A y B, ambas n × n, se llaman semejantes cuando representan una misma transformación lineal referida a bases distintas.*

La relación de semejanza puede formularse sin necesidad de conceptos distintos al de matriz, por una exposición semejante a la de las transformaciones conjugadas, desarrollada en Cap. VI, § 11. Se observará que cualquier cambio de base en un espacio vectorial de n dimensiones $V_n(F)$ reemplaza las antiguas coordenadas X de un vector ξ por nuevas coordenadas $X^* = XP$ para alguna matriz regular P (cfr. Cap. IX, § 4). En las antiguas ecuaciones $Y = XA$ de la transformación dada T, podemos reemplazar a la par $X = X^*P^{-1}$ e $Y = Y^*P^{-1}$ por las nuevas coordenadas, resultando

(24) $$Y^*P^{-1} = Y = XA = X^*P^{-1}A, \qquad Y^* = X^*(P^{-1}AP).$$

Visualicemos esta ecuación así: $X^* \xleftarrow{P} X \xrightarrow{A} Y \xrightarrow{P} Y^*$. Para ir desde X^* hasta Y^* vamos primero a X por P^{-1}, luego iremos de X a Y por A, y por fin de Y a Y^* por P. La conclusión de (24) es:

TEOREMA 25. *Dos matrices n × n son semejantes si, y sólo si, $B = P^{-1}AP$ para alguna matriz regular P.*

Resulta como corolario que dos matrices semejantes tienen el mismo determinante; además, son equivalentes en el sentido de § 4.

El álgebra de matrices se aplica con especial comodidad a las matrices diagonales; para sumar o multiplicar dos de ellas, basta sumar o multiplicar los correspondientes elementos diagonales. Por esta y otras razones, es interesante saber cuándo una matriz es se-

mejante a una matriz diagonal, y también cuándo dos matrices diagonales son equivalentes entre sí. La resolución de estas cuestiones envuelve las nociones de vectores característicos y de raíces características, llamados también, en la mecánica cuántica, estados propios y valores propios (o autovalores).

Para ilustrar con un ejemplo la investigación de una matriz diagonal semejante a una matriz dada, consideremos la transformación $y_1 = -x_1 - ax_2$, $y_2 = x_2$ (un corrimiento de cizalla sobre el eje x_1 seguido por una reflexión en el eje x_2). Nos preguntamos cuándo el vector transformado (y_1, y_2) es igual al producto $c(x_1, x_2)$ del vector original (x_1, x_2) por un escalar c.

Esto lleva a las ecuaciones

$$y_1 = -x_1 - ax_2 = cx_1, \qquad y_2 = x_2 = cx_2.$$

Si $x_2 = 0$, será $c = -1$, y el vector será $(x_1, 0)$. Si $x_2 \neq 0$, será $c = 1$ y $2x_1 = -ax_2$, luego el vector será $(x_1, -2x_1/a)$. En ambos vectores $(x_1, 0)$ y $(x_1, -2x_1/a)$ el escalar x_1 puede tomarse arbitrariamente; en particular, resulta que los vectores $\epsilon_1^* = (1, 0)$ y $\epsilon_2^* = (a, -2)$ son transportados por T sobre sus múltiplos escalares, $\epsilon_1^* T = -\epsilon_1^*$, $\epsilon_2^* T = \epsilon_2^*$ (ilústrese esto con un diagrama que muestre el efecto sobre ϵ_1^* y ϵ_2^* del corrimiento y reflexión T, para $a = 3$). Estos vectores ϵ_1^*, ϵ_2^* son independientes, y por lo tanto proporcionan los ejes (oblicuos) de un nuevo sistema de coordenadas. Si las nuevas coordenadas del vector ξ son x_1^* y x_2^*, será

$$\xi T = (x_1^* \epsilon_1^* + x_2^* \epsilon_2^*) T = x_1^* (\epsilon_1^* T) + x_2^* (\epsilon_2^* T) = -x_1^* \epsilon_1^* + x_2^* \epsilon_2^*$$

Por lo tanto, las ecuaciones de la transformación tienen la forma sencilla $y_1^* = -x_1^*$, $y_2^* = x_2^*$; deducimos que T es una «simetría oblicua» según la dirección de ϵ_1^* respecto al nuevo eje ϵ_2^*.

Iniciaremos un análisis semejante para una transformación lineal arbitraria T, definiendo como *vector característico* de T a todo vector ξ no nulo tal, que $\xi T = c\xi$ para algún escalar c; mientras que llamaremos *raíces características* de T a aquellos escalares c tales, que $\xi T = c\xi$ para algún vector ξ no nulo.

Las matrices semejantes corresponden a una misma transformación lineal, y por lo tanto tienen las mismas raíces características.

Si D es una matriz diagonal, y son $d_1, \ldots, d_n$ sus elementos diagonales, los vectores unidad $\epsilon_1 = (1, 0, \ldots, 0), \ldots, \epsilon_n = (0, \ldots, 0, 1)$

serán vectores característicos para D, ya que $\epsilon_1 D = d_1\epsilon_1, \ldots, \epsilon_n D = d_n\epsilon_n$. Vemos también que los elementos de la diagonal principal son las raíces características. Recíprocamente, supongamos que los vectores característicos de una matriz A sean suficientes para engendrar todo el espacio $V_n(F)$ sobre el que A opera. Entonces (Capítulo VII, Teor. 8, Cor. 2) podremos extraer un subconjunto $\beta_1, \ldots, \beta_n$ de vectores característicos, que constituyen una base de $V_n(F)$. Como $\beta_1 A = c_1\beta_1, \ldots, \beta_n A = c_n\beta_n$, al referir a esta nueva base la transformación lineal determinada por A, vendrá representada por una matriz diagonal C, siendo $c_1, \ldots, c_n$ sus elementos diagonales. Así hemos demostrado

TEOREMA 26. *Una matriz* n × n *es semejante a una matriz diagonal si, y sólo si, sus vectores característicos engendran el* $V_n(F)$.

Para hallar explícitamente los vectores característicos de tal matriz A, observemos que X es uno de ellos si, y sólo si, $XA = \lambda X$ para algún *escalar* $\lambda = c$. Pero esto es lo mismo que decir que $X(A - \lambda I) = 0$ para algún escalar λ. Para un valor fijo de λ, la determinación del conjunto de tales vectores equivale a resolver un sistema de ecuaciones lineales homogéneas, lo que puede hacerse por métodos elementales (Cap. II, § 8).

Por lo tanto, deberemos hallar primero las raíces características, o sea : los valores de λ para los cuales el sistema homogéneo $X(A - \lambda I) = 0$ tiene solución $X \neq 0$. Pero, por el Teorema 17 y el corolario del Teorema 8, esto sucede si $|A - \lambda I| = 0$, y sólo en este caso. Pero como un determinante es un polinomio lineal en los elementos de cada fila, $|A - \lambda I|$ es un polinomio de grado n en λ de la forma

$$(25) \qquad |A - \lambda I| = (-1)^n\lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0.$$

Si la matriz A tiene sus elementos reales o complejos, la ecuación tiene al menos una raíz compleja, por el Teorema fundamental del álgebra, luego : *una matriz compleja tiene por lo menos un vector característico*.

El polinomio (25) es llamado comúnmente *polinomio característico* de A. El razonamiento precedente sobre (25) demuestra que

TEOREMA 27. *Las raíces características de una matriz* n × n A, *son las raíces del polinomio característico de* A.

Ejemplo. Busquemos la matriz diagonal semejante a la $\begin{pmatrix} -3 & 4 \\ 2 & -1 \end{pmatrix}$. Calculando el polinomio característico de esta matriz resulta $\lambda^2+4\lambda-5$. Sus raíces son 1 y 5; luego los vectores característicos satisfacen a uno o a otro de los sistemas homogéneos que siguen:

$$\begin{aligned} -3x+2y&=x \qquad & -3x+2y&=-5x \\ 4x-y&=y \qquad & 4x-y&=-5y \end{aligned}$$

Resolviéndolos, se encuentran los vectores característicos (1, 2) y (1, —1). Utilizándolos como nueva base, la transformación tomará forma diagonal. La nueva matriz diagonal se expresa, de acuerdo con el Teorema 25, como un producto

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -3 & 4 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -5 \end{pmatrix}.$$

Algunas veces se pueden hallar formas canónicas para toda una familia de transformaciones lineales. Por ejemplo, consideremos las transformaciones lineales de período dos, es decir, tales, que $T^2=I$. Para concretar, supongamos que T opera sobre $V_n(F)$ y que en el campo base es $1+1\neq 0$. Entre los vectores característicos para T se incluyen todos los vectores no nulos $\eta=\xi(T+I)$ del resultante de $(T+I)$, puesto que

$$[\xi(T+I)]T=\xi(T^2+T)=\xi(T+I).$$

También se incluyen todos los vectores no nulos del resultante de $(T-I)$, ya que

$$[\xi(T-I)]T=\xi(T^2-T)=\xi(I-T)=-[\xi(T-I)].$$

Pero como $1+1\neq 0$, cualquier ξ puede escribirse como una suma

$$\xi=(1/2)[\xi(T+I)-\xi(T-I)];$$

luego los vectores característicos con raíces características ± 1, engendran todo el espacio. Por lo tanto, T puede representarse por una, al menos, de las siguientes matrices diagonales:

$$I=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\quad A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\quad B=\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$-I=\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

El método de reducir una matriz a forma diagonal por descomposición factorial de una ecuación polinómica a la que satisface, puede generalizarse :

TEOREMA 28. *Una matriz* A *es semejante a una matriz diagonal si, y sólo si,* A *satisface a una ecuación polinómica que sea producto de factores lineales distintos.*

Demostración. Primero, supongamos que A satisface a una ecuación polinómica mónica, que puede ser descompuesta a factores lineales distintos,

$$(26) \qquad (A-\lambda_1 I)(A-\lambda_2 I)\dots(A-\lambda_s I)=p(A)=0.$$

Designemos por $q_k(A)$ el producto de los $(A-\lambda_j I)$ con $j\neq k$; así tendremos s polinomios de grado $(s-1)$. Cualquier vector $\eta=\xi q_k(A)$ en el *resultante* de un $q_k(A)$ será característico, pues se tiene

$$(27) \qquad \eta A=\lambda_k\eta+\xi[q_k(A)\cdot A-\lambda_k q_k(A)]=\lambda_k\eta+\xi p(A)=\lambda_k\eta+0=\lambda_k\eta.$$

Si, por consiguiente, los resultantes de los diferentes $q_k(A)$ engendran la totalidad del espacio, por el Teorema 26, A será semejante a una matriz diagonal. Por otra parte, la hipótesis de que los factores lineales de (26) sean todos distintos implica que los polinomios $q_k(A)$ no tengan factor común en A. El máximo común divisor I de estos polinomios puede entonces escribirse como en Capítulo IV, en la forma

$$I=f_1(A)q_1(A)+\dots+f_s(A)q_s(A),$$

donde las f_i indican también polinomios en A. Esto significa que cualquier vector ξ puede escribirse como

$$\xi=\xi I=\sum_k \xi f_k(A)q_k(A)=\eta_1+\dots+\eta_s,$$

donde cada vector $\eta_k=\xi[f_k(A)q_k(A)]$ está en el resultante de $q_k(A)$, y por lo tanto es un vector característico con $\eta_k A=\lambda_k\eta_k$, como en (27). Por lo tanto, A es semejante a una matriz diagonal.

Recíprocamente, si D es una matriz diagonal cuyos diferentes elementos diagonales son $c_1, \dots, c_s$, la transformación determinada

por el producto $q(D)=(D-c_1I)\dots(D-c_sI)$ transforma todos los vectores de la base en cero, y por tanto, $q(D)=0$. Luego si $A=P^{-1}DP$ es semejante a D, será

$$0=P^{-1}(D-c_1I)\dots(D-c_sI)P=\prod_{k=1}^{s}[P^{-1}(D-c_kI)]P=$$

$$=\prod_{k=1}^{s}(P^{-1}DP-c_kP^{-1}IP)=\prod_{k=1}^{s}(A-c_kI)=q(A).$$

Por lo tanto, A satisface a una ecuación polinómica de la forma enunciada.

En el Cap. VIII, § 6, definimos la ecuación mínima de A como la de grado mínimo entre todas las ecuaciones mónicas que tienen la solución A. Por lo tanto, la conclusión del Teorema 28 se puede enunciar así: A *es semejante a una matriz diagonal si, y sólo si, el primer miembro de la ecuación mínima de* A *es igual a un producto de factores lineales distintos.*

Apéndice. La igualdad $P^{-1}q(D)P=q(P^{-1}DP)$, utilizada antes, es cierta en general. Para demostrarlo, notemos primero el siguiente resultado:

TEOREMA 29. *Para cualquier matriz* n × n *regular* P, *la correspondencia* $A\to P^{-1}AP$ *es un automorfismo del álgebra de las matrices* n × n *(sobre el campo de escalares).*

Pues $P^{-1}(A+B)P=P^{-1}AP+P^{-1}BP$ por la ley distributiva, $P^{-1}(AB)P=(P^{-1}AP)(P^{-1}BP)$ y $P^{-1}(cA)P=c(P^{-1}AP)$.

Como corolario resulta que si A y B, ambas $n\times n$, son semejantes, serán equivalentes en el grupo de todos los automorfismos del álgebra de las matrices $M_n(F)$; además, si dos matrices regulares son semejantes, serán *conjugadas* en el grupo de las matrices regulares (el grupo lineal). Finalmente, también implica el teorema que las matrices semejantes tengan la misma ecuación mínima.

EJERCICIOS

1. Demostrar que las ecuaciones $2x'=(1+b)x+(1-b)y$, $2y'=(1-b)x+(1+b)y$ representan una compresión contra la recta que pasa por el origen inclinada 45°. Calcular las raíces características y los vectores característicos de la transformación e interpretarlo todo geométricamente.

2. Calcular las raíces características y los vectores característicos de las matrices

a) $\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$; b) $\begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$; c) $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$; d) $\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$;

e) $\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}$; f) $\begin{pmatrix} 4 & 9 & 0 \\ 0 & -2 & 8 \\ 0 & 0 & 7 \end{pmatrix}$.

3. Para cada matriz A del Ejerc. 2 hallar, cuando sea posible, una matriz regular P para la que PAP^{-1} sea diagonal.

4. a) Hallar las raíces características complejas de la matriz que representa una rotación del plano de amplitud θ.
   b) Demostrar que la matriz que representa una rotación del plano de ángulo θ $(0 < \theta < \pi)$ no es semejante a ninguna matriz diagonal real.

5. Demostrar que ninguna matriz $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ es semejante a una matriz diagonal real o compleja. Interpretarlo geométricamente.

6. a) Demostrar que cualquier matriz real 2×2 satisfaciendo a $A^2 = -I$ es semejante a la matriz $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.
   b) Demostrar que ninguna matriz real 3×3 satisface a $A^2 = -I$.
   c) ¿Qué puede decirse de las matrices A reales 4×4, satisfaciendo a $A^2 = -I$?

7. a) Demostrar que el resultante y el espacio nulo de cualquier transformación lineal «idempotente» T satisfaciendo a $T^2 = T$, son espacios complementarios.
   b) Demostrar que dos matrices idempotentes que tengan la misma característica, son semejantes. [*Sugerencia:* Utilizar el resultado de a).]

8. a) Clasificar todas las matrices 3×3 complejas que satisfacen a $A^3 = I$.
   b) Hacer lo mismo para las matrices 3×3 reales.

9. a) Demostrar que cualquier matriz 2×2 tal, que $X^2 = 0$, es semejante a $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ o es $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
   b) Demostrar el resultado análogo para matrices 3×3.

10. Cualquier corrimiento plano satisface a $A^2 + I = A + A$. Hallar la forma canónica de las matrices 2×2 que satisfacen a esta ecuación. (*Sugerencia:* Formar $A - I$.)

11. Demostrar que el conjunto de todos los vectores característicos que corresponden a una raíz característica determinada de una matriz dada, constituyen un subespacio lineal, cuando O se incluye entre los vectores característicos.

12. Demostrar que cualquier matriz real 2×2 cuyo determinante sea negativo, es semejante a una matriz diagonal. (Interpretarlo geométricamente.)

13. a) ¿Es el polinomio $A^3 - I$ el producto de factores lineales distintos sobre el campo de enteros mód. 2?
   * b) ¿Bajo qué condiciones para el campo de escalares, una matriz A que satisface a $A^2 + A = 2I$, es semejante a una matriz diagonal?

[illegible]. Demostrar que si el polinomio característico de una matriz compleja A [illegible] tiene raíces múltiples, entonces A es semejante a una matriz diagonal.

[illegible]. Demostrar que toda matriz $(2n+1)\times(2n+1)$ real tiene un vector carac[illegible] real.

[illegible]. Polinomio característico de una matriz

[illegible] noción de polinomio característico de una matriz ha sido [illegible]da en § 8; en esta sección será aplicado de diversos modos. [illegible] demostraremos que dos matrices diagonales semejantes tienen [illegible] los elementos diagonales (tal vez permutados), y aplicare[illegible] a probar que los coeficientes d_i de una forma diagonal cua[illegible] $d_1x_1^2+\dots+d_nx_n^2$ son invariantes en el grupo ortogonal (Ca[illegible] [illegible], § 9). Veremos que cualquier matriz A de tipo $n\times n$ satis[illegible] una ecuación polinómica de grado n (y no precisamente de [illegible] n^2, como se estableció en Cap. VIII, § 8).

[illegible] coeficientes de esta ecuación característica son *invariantes* [illegible] grupo $A \to P^{-1}AP$ (P es regular); estos coeficientes son poli[illegible] formados con los elementos de A, y constituyen, precisa[illegible], un sistema completo de invariantes para este grupo.

DEFINICIÓN. *Sea A una matriz $n\times n$, con coeficientes en un [illegible] F. El determinante*

$$(1)\qquad |A-\lambda I| = b_0 + b_1\lambda + \dots + b_{n-1}\lambda^{n-1} + (-1)^n\lambda^n$$

[illegible] polinomio característico de A.

[illegible] interpretar justamente [illegible] definición, debemos consi[illegible] λ como una indetermi[illegible] que la matriz $A-\lambda I$ [illegible] los coeficientes en el do[illegible] $F[\lambda]$ de las formas po[illegible] en λ sobre F. Entre [illegible]ficientes (invariantes) de $|A-\lambda I|$ merecen especial men[illegible] $b_0 = |A|$ y $b_{n-1} = \pm(a_{11}+\dots+a_{nn})$.

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix}$$

[illegible] 10. *Las matrices semejantes tienen el mismo poli[illegible] característico.*

Demostración. Sean las matrices A y $B = P^{-1}AP$. Como $|P^{-1}| = |P|^{-1}$ y $|P|$ son escalares, pueden conmutarse, y la regla para multiplicar determinantes dará

$$|P^{-1}AP - \lambda I| = |P^{-1}AP - \lambda P^{-1}IP| = |P^{-1}(A - \lambda I)P| =$$
$$= |P^{-1}| \cdot |A - \lambda I| \cdot |P| = |A - \lambda I|.$$

Teorema 31. *El polinomio característico de una matriz diagonal* D, *con elementos diagonales* $d_1, \ldots, d_n$, *es*

$$|D - \lambda I| = (d_1 - \lambda)(d_2 - \lambda) \ldots (d_n - \lambda).$$

La demostración es inmediata, ya que $D - \lambda I$ es asimismo una matriz diagonal. Resulta, como corolario, que el conjunto de los elementos diagonales coincide con el de las raíces del polinomio característico (y con el mismo orden de multiplicidad). Por lo tanto, el conjunto de elementos diagonales y el de repeticiones en cada diagonal son los mismos en dos matrices diagonales semejantes. Lo cual puede formularse así:

Corolario. *Dos matrices diagonales son semejantes si y sólo si, difieren únicamente en el orden de sus términos diagonales.*

Las propiedades de semejanza lanzan nueva luz sobre las transformaciones ortogonales de una forma cuadrática real [Cap. IX, § 9]. Si una forma cuadrática XAX^T con matriz A puede reducirse por una transformación ortogonal $Z = XP$ a una forma diagonal $\lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$, la matriz diagonal de la nueva forma es $D = PAP^T$. Como P es ortogonal, $P^T = P^{-1}$ y $D = PAP^{-1} = (P^{-1})^{-1}A(P^{-1})$; luego A y D son matrices semejantes. Las raíces características $\lambda_1, \ldots, \lambda_n$ de D son, pues, las raíces características de la matriz dada A. Esto da al Teorema 19 del Cap. IX la penetrante forma que sigue:

Teorema 32. *Cualquier forma cuadrática* XAX^T *puede reducirse por una transformación ortogonal de variables a una forma diagonal* $\lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$, *en la que los coeficientes* λ_i *son las raíces del polinomio característico* $|A - \lambda I| = (\lambda_1 - \lambda) \ldots (\lambda_n - \lambda)$ *de* A.

Pero la ecuación característica, y por ende sus raíces, está determinada unívocamente por A. Esto prueba la unicidad esencial de la forma diagonal, y da una manera directa de calcular sus coeficientes. Se pueden también hallar directamente los ejes principa-

por el método de este Capítulo, sin acudir a los razonamientos máximos hechos en el Cap. IX (ver el siguiente ejercicio 4).

Como ya sabemos que cualquier matriz simétrica real es equi-nte, ortogonalmente, a una matriz diagonal real, obtenemos

COROLARIO. *Todas las raíces características de una matriz real imétrica son reales.*

Los vectores característicos proporcionan un método efectivo la construcción de la matriz ortogonal P que efectúe la reduc- de XAX' a forma diagonal, por lo menos en el caso de raíces cterísticas distintas. Por ejemplo, sea A una matriz 2×2, con es características $\lambda_1 \neq \lambda_2$, y vectores característicos correspon-tes X_1 y X_2. La expresión bilineal $X_1 A X_2'$ puede entonces cal-rse de dos modos distintos como

$$(X_1 A)X_2' = \lambda_1 (X_1 X_2'), \qquad X_1(AX_2') = X_1(X_2 A)' = \lambda_2 (X_1 X_2').$$

mo $\lambda_1 \neq \lambda_2$, $X_1 X_2'$ será nulo, y X_1 será ortogonal a X_2. Esco-do convenientemente los escalares c_1 y c_2, la matriz P de filas X_1 y $c_2 X_2$ será ortogonal. Entonces el cálculo demuestra que P^{-1} es una matriz diagonal; por lo tanto, P es la matriz de transformación ortogonal que reduce XAX' a forma diagonal.

TEOREMA 33 (Cayley-Hamilton). *Cualquier matriz cuadrada sa-ace a su ecuación característica.*

Esto quiere decir que, si cada potencia λ^i en el polinomio carac-ístico $f(\lambda) = |A - \lambda I|$ de (28) se reemplaza por la misma poten- A^i de la matriz (y si λ^0 se reemplaza por $A^0 = I$), el resultado cero;

$$b_0 I + b_1 A + \ldots + b_{n-1} A^{n-1} + (-1)^n A^n = 0.$$

Demostración. Los elementos de la matriz $A - \lambda I$ son poli-ios lineales en λ, luego los determinantes de sus menores serán inomios en λ de grado $n-1$ a lo más. Cada elemento de la ma- C adjunta de $A - \lambda I$ es uno de tales menores, así que esta unta puede escribirse como una suma de n matrices, en las que recen términos con una potencia determinada λ^0, λ^1, ..., λ^{n-1} λ. Es decir, la adjunta $C = C(\lambda)$ es un polinomio $C = C(\lambda) = \Sigma C_i \lambda^i$

cuyos coeficientes son matrices C_i. De acuerdo con (20), el producto de $A-\lambda I$ por su adjunto es

$$C(\lambda)(A-\lambda I)=|A-\lambda I|\,I=f(\lambda)I, \tag{30}$$

donde $f(\lambda)$ es el polinomio característico. Podemos considerar $f(\lambda)I=f(\lambda I)$ como otro polinomio en λ cuyos coeficientes son matrices. La ecuación (30) nos dice entonces que este polinomio da residuo cero cuando se le divide *a la derecha* por el polinomio lineal $(A-\lambda I)$. Si el teorema del resto para polinomios ordinarios (Capítulo IV, Teor. 8, Cor. 1) se aplicase a este caso, nos diría que el resto de la división de $f(\lambda I)$ por $A-\lambda I$ es igual a $f(A)$ y, por lo tanto, sería $f(A)=0$, c. q. d. Por lo tanto, sólo nos hace falta demostrar que el teorema del resto es aplicable aunque los coeficientes de nuestros «polinomios» no sean permutables entre sí.

A este fin, observemos que la conocida descomposición factorial

$$A^i-\lambda^i I=(A^{i-1}+\lambda A^{i-2}+\ldots+\lambda^{i-1}I)(A-\lambda I)$$

dará, en función de los coeficientes b_i de (28),

$$f(A)-f(\lambda I)=\sum_{i=0}^{n} b_i A^i-\sum_{i=0}^{n} b_i\lambda^i I=\sum_{i=0}^{n} b_i(A^i-\lambda^i I)=$$

$$=\sum_{i=0}^{n} b_i(A^{i-1}+\lambda A^{i-2}+\ldots+\lambda^{i-1}I)(A-\lambda I),$$

$$f(A)-f(\lambda I)=-G(\lambda)(A-\lambda I), \tag{31}$$

donde $G(\lambda)$ es un nuevo polinomio con matrices en los coeficientes. Esta ecuación muestra que el resto de la división de $f(\lambda I)$ por $A-\lambda I$ es $f(A)$. Probemos ahora que el resto es único.

Si sumamos (31) y (30) obtendremos

$$f(A)=[C(\lambda)-G(\lambda)](A-\lambda I).$$

Si $C(\lambda)\neq G(\lambda)$, la expresión entre corchetes es un polinomio matricial en λ de grado 1 por lo menos; pero en el primer miembro de la igualdad no interviene λ. Esta contradicción puede explicarse tan sólo si es $C(\lambda)=G(\lambda)$, en cuyo caso $f(A)=0$, c. q. d.

Teorema 34. *Si una matriz* $n\times n$ *tiene su polinomio característico* $f(\lambda)$ *con* n *raíces distintas, es semejante a una matriz diagonal cuyos elementos diagonales son tales raíces.*

Demostración. Sea A la matriz $n \times n$. Como $f(A)=0$, A será, por el Teor. 28, semejante a una matriz diagonal D. Como A y D tienen el mismo polinomio característico (Teorema 30), los elementos diagonales de D son también, por el Teor. 31, las raíces de $f(\lambda)$.

Una matriz simétrica real es semejante a una matriz diagonal aunque las raíces características no sean distintas, pero el mismo resultado no es válido, en general, para las matrices complejas (cuyos polinomios característicos pueden descomponerse siempre en factores lineales). Por ejemplo, la matriz triangular

$$(32) \qquad A=\begin{pmatrix} c & 1 \\ 0 & c \end{pmatrix}, \qquad |A-\lambda I| = \begin{vmatrix} c-\lambda & 1 \\ 0 & c-\lambda \end{vmatrix} = (\lambda - c)^2$$

tiene sólo una raíz característica c. Los únicos vectores característicos son los $(0, b)$, todos ellos múltiplos del vector característico $(0, 1)$; por lo tanto (Teorema 26), A no es semejante a una matriz diagonal.

En la teoría general de las formas canónicas de una transformación lineal con raíces características múltiples, aparecen como formas canónicas, matrices cuya diagonal está formada por bloques de los tipos siguientes:

$$(33) \qquad B=\begin{bmatrix} \lambda_0 & 1 & 0 & 0 \\ 0 & \lambda_0 & 1 & 0 \\ 0 & 0 & \lambda_0 & 1 \\ 0 & 0 & 0 & \lambda_0 \end{bmatrix} \qquad C_f=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -c_0 & -c_1 & -c_2 & -c_3 \end{bmatrix}$$

El polinomio característico del primero es $(\lambda - \lambda_0)^4$, y el del segundo $f(\lambda)=\lambda_4+c_3\lambda^3+c_2\lambda^2+c_1\lambda+c_0$; la segunda matriz C_f es llamada a veces *matriz asociada* a la forma polinómica $f(\lambda)$. Los invariantes de la matriz en su reducción a tales bloques son los llamados «divisores elementales» de la misma.

EJERCICIOS

1. Sea D una matriz diagonal con los elementos $(3, 1, -1)$ y sea P una matriz triangular con las filas $(1, 2, -3)$, $(0, -1, 4)$, $(0, 0, 1)$. Calcular la ecuación característica de $P^{-1}DP$ y compararla con la de D.
2. Demostrar que si A es triangular (es decir, si $a_{ij}=0$ para $i > j$), el polinomio característico de A es $(a_{11}-\lambda)\ldots(a_{nn}-\lambda)$.
3. a) Si una matriz $n \times n$, A, tiene raíces características distintas, mostrar como los vectores característicos pueden utilizarse para construir efectivamente una matriz P para la cual PAP^{-1} sea diagonal.
   b) Demostrar el Teor. 34 sin utilizar el Teorema de Cayley-Hamilton.

4. En Ejerc. 3, mostrar como P puede hacerse ortogonal si A es real y simétrica.
5. Demostrar que si A es simétrica, $\xi A = a\xi$ y $\eta A = b\eta$ $(a \neq b)$, entonces $\xi \perp \eta$.
6. Escribir una forma diagonal cuadrática equivalente en las transformaciones ortogonales, a las formas dadas. (*Sugerencia:* En b), mostrar que todas las raíces características son múltiplos de 9.)
   a) $x^2+6xy-2y^2-2yz+z^2$,
   b) $-2x^2-11y^2-5z^2+4xy+16yz+20xz$,
   c) $3x^2-y^2-3z^2-t^2-4xz-10yt$.
7. Presentar una transformación ortogonal que reduzca cada forma del Ejerc. 6 a su diagonal equivalente.
8. Hallar condiciones necesarias y suficientes para que las raíces características de una matriz 2×2 sean iguales.
9. Hallar todas las matrices 2×2 con raíces características $+1$ y -1.
10. Demostrar, por sustitución directa, que cualquier matriz 2×2 satisface a su ecuación característica.
11. Demostrar que las matrices transpuestas A y A' tienen la misma ecuación característica.
12. Demostrar que, en (28), $b_{n-1} = \pm(a_{11}+\ldots+a_{nn})$ (el invariante $a_{11}+\ldots+a_{nn}$ se llama la *traza* de A).
13. Demostrar directamente, por la definición, que todas las raíces características de una matriz A real y simétrica, son reales. (*Sugerencia:* Para un X característico, mostrar $XAX^{*\prime} = \lambda XX^{*\prime} = \lambda^* X^* X'$, designando X^* el complejo conjugado de X.)
14. a) Demostrar que las raíces características de una matriz hermítica son reales.
    b) Demostrar que los vectores característicos engendran el espacio de todos los vectores.
*15. Si λ es una raíz característica de una matriz unitaria U, demostrar que $|\lambda| = 1$.
*16. Si Q y R son dos matrices ortogonales y ambas reducen a una misma matriz simétrica A a las formas diagonales $Q'AQ$ y $R'AR$, y si A es regular y tiene raíces características distintas, demostrar que $R = QP$, donde P es una matriz de permutación.
*17. Demostrar el Teorema de los ejes principales para una matriz A real y simétrica, por el siguiente análisis de la transformación lineal $X \to XA$:
   a) La matriz A tiene un vector característico α_1 de longitud 1.
   b) Si α_1 se toma como primer vector en una nueva base ortonormal, la nueva matriz de la transformación dada tiene ceros en la primera columna y en la primera fila, excepto para el primer elemento.
   c) La demostración se concluye por inducción.
18. a) Calcular los polinomios característicos para la matriz B presentada como forma canónica en (33).
    b) Mostrar explícitamente que B satisface a su ecuación característica.
19. Lo mismo para la matriz 4×4 C_f, asociada de (33).
20. Mostrar que la matriz 4×4 asociada C_f no satisface a ninguna ecuación polinómica de grado menor de 4.

*21. a) Escribir inmediatamente la matriz asociada a cualquier forma polinómica mónica de grado n, y calcular su polinomio característico.
b) Demostrar directamente que una matriz asociada cumple siempre su ecuación característica.

22. Para una matriz dada A, consideremos el conjunto de todos los polinomios $g(\lambda)$ con $g(A)=0$.
a) Mostrar que este conjunto es cerrado para la adición, sustracción y multiplicación por cualquier polinomio.
b) Mostrar que cualquier polinomio del conjunto $g(\lambda)$ es múltiplo del polinomio mínimo de A.
c) Demostrar que el polinomio característico de cualquier matriz es un múltiplo de su polinomio mínimo.

*23. Demostrar que el volumen del elipsoide $\Sigma a_{ij}x_i x_j < 1$ es $4\pi/3\sqrt{|A|}$, con $A=\| a_{ij} \|$. (*Sugerencia:* Referirlo a los ejes principales y utilizar el Teorema 24.)

CAPÍTULO XI

Algebra de clases

1. Definiciones fundamentales

El concepto de «clase de objetos» es fundamental en la lógica; no es posible conducir un razonamiento lógico sin implicarlo. Es tan fundamental y familiar dicho concepto, que no vale intentar definirlo mediante otros conceptos básicos, ni ejemplarizarlo con ilustraciones triviales. En vez de esto, haremos mención de ideas abstractas sinónimas, como las de «conjunto de elementos» o «colección de cosas» o «agregación de individuos».

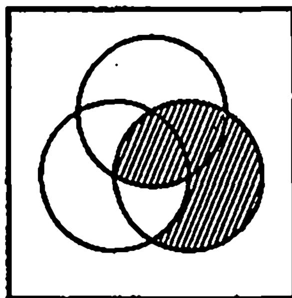


Figura 1

Por otra parte, poca gente sabe que de este concepto abstracto se deduce un álgebra de conjuntos. Esto no es extraño, pues la importancia de esta álgebra ha sido ignorada hasta bien recientemente, incluso por los propios matemáticos. El presente capítulo será dedicado a presentar el álgebra de clases (o álgebra booleana) y sus generalizaciones (teoría de las redes).

Comenzaremos por definir las relaciones y operaciones del álgebra de conjuntos; para esto, supondremos en lo que sigue que I es conjunto y que $X, Y, Z, \ldots$ designan subconjuntos de I. Así, I puede ser el conjunto de los puntos de un cuadro, y X, Y, Z, los de diversas regiones en el interior de I (ver fig. 1).

Convendremos en escribir $X \leqslant Y$ [léase «X está contenido (o incluído) en Y»] cuando cada elemento de X esté en Y.

Escribiremos $X \frown Y$ (léase «la intersección de X e Y») para indicar el conjunto de todos los elementos simultáneamente en ambos X e Y, y escribiremos $X \smile Y$ (léase «la reunión de X e Y») para designar el conjunto de los elementos en cualquiera X o Y, o en ambos. Los símbolos $\frown$ y $\smile$ se llamarán «cap» y «cup» respectivamente, palabras arbitrarias que introducimos para aligerar la lectura de las fórmulas.

Finalmente; escribiremos X' (y diremos «complemento de X») para designar el conjunto de todos los elementos que no están en X. El signo $'$ se llamará «prima», como es habitual.

Las operaciones del álgebra de conjuntos o clases pueden ser ilustradas gráficamente por medio de los diagramas de Venn. Así, en la fig. 1, X, Y y Z aparecen como los interiores de tres círculos iguales interiores al cuadrado I. Las combinaciones de estos espacios pueden distinguirse por medio de sombras diversas: Y' es el exterior de Y, y $X \frown (Y' \smile Z)$ es el área que aparece rayada.

EJERCICIOS

1. El diagrama de Venn para X, Y y Z separa el cuadrado en ocho áreas no rampantes. Encontrar para cada una de ellas una combinación algebraica entre X, Y y Z que la represente exactamente.
2. Sombrear sobre el diagrama de Venn cada una de las siguientes áreas:

$$(X' \frown Y) \smile (X \frown Z'), \quad (X \smile Y)' \frown Z, \quad (X \smile Y') \smile Z'.$$

3. Por el sombreado de las apropiadas áreas del diagrama de Venn determinar cuáles de las siguientes igualdades son válidas:

a) $(X' \smile Y)' = X \frown Y'$;
b) $X' \smile Y' = (X \smile Y)'$;
c) $(X \smile Y) \frown Z = (X \smile Z) \frown Y$;
d) $X \smile (Y \frown Z)' = (X \smile Y') \frown Z'$.

2. Leyes: Analogía con la Aritmética

Las relaciones y operaciones que acabamos de definir gozan de varias propiedades algebraicas, muchas de las cuales son idénticas a las conocidas leyes de Aritmética. Esta es la razón de que el Álgebra de clases haya venido a ser un instrumento práctico para el cálculo.

Así, la relación de inclusión comparte con los ordinarios «menor o igual que» los siguientes caracteres:

Reflexivo: Para todo X, $X \leqslant X$;
Antisimétrico: Si $X \leqslant Y$ e $Y \leqslant X$, resulta $X = Y$;
Transitivo: Si $X \leqslant Y$ e $Y \leqslant Z$, resulta $X \leqslant Z$.

Las operaciones de intersección y unión tienen cierta analogía con la adición y multiplicación ordinarias, como se desprende de las siguientes leyes:

Idempotente: $X \frown X = X$ y $X \smile X = X$;
Conmutativa: $X \frown Y = Y \frown X$ y $X \smile Y = Y \smile X$;
Asociativa: $X \frown (Y \frown Z) = (X \frown Y) \frown Z$ y
$X \smile (Y \smile Z) = (X \smile Y) \smile Z$;
Distributiva: $X \frown (Y \smile Z) = (X \frown Y) \smile (X \frown Z)$ y
$X \smile (Y \frown Z) = (X \smile Y) \frown (X \smile Z)$.

Es claro que, excepto la ley de idempotencia y la segunda ley distributiva, las leyes que acabamos de formular se corresponden con las de la Aritmética ordinaria si reemplazamos «cup» y «cap» por «más» y «por» respectivamente.

Las operaciones de intersección y reunión están relacionadas con la inclusión por un principio fundamental de

Conformidad: Las tres condiciones $X \leqslant Y$, $X \frown Y = X$ y $X \smile Y = Y$ son equivalentes entre sí.

Designemos el conjunto vacío o nulo por 0; entonces, para los conjuntos 0 e I podrán formularse las siguientes propiedades:

Cotas universales: $0 \leqslant X \leqslant I$ para todo X;
Intersección: $0 \frown X = 0$ e $I \frown X = X$;
Reunión: $0 \smile X = X$ e $I \smile X = I$.

Las propiedades de intersección y reunión son, excepto la última, análogas a las propiedades del cero y del uno en la Aritmética ordinaria.

Finalmente, la operación de complementar está relacionada con las de intersección y reunión por tres nuevas leyes:

Complementación: $X \cap X' = 0$ y $X \cup X' = I$;
Dualidad: $(X \cap Y)' = X' \cup Y'$ y $(X \cup Y)' = X' \cap Y'$;
Involución: $(X')' = X$.

La primera y última de estas leyes se aplica también a la Aritmética booleana, si X' se reemplaza por $1 - X$.

Debe observarse que hemos establecido estas leyes, pero no las hemos probado. La causa de esto es su carácter fundamental. En efecto, hay paráfrasis de los principios de lógica universalmente aceptados, que van implicados en cada parte de la argumentación deductiva. Si se comienza el estudio de la lógica por la lógica de clases, estos principios aparecerán esencialmente como postulados. Otras analogías son posibles, como vamos a ver.

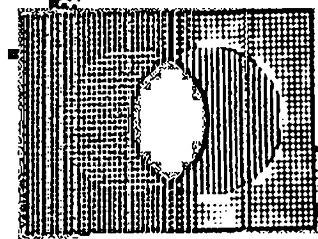


Figura 2

Primeramente, podemos inducir estas leyes partiendo de ejemplos particulares. Tales son los que ofrece un apropiado diagrama de Venn. Si X e Y son, respectivamente, el interior del círculo de la izquierda y del de la derecha de la figura 2, se observa que X' es el área sombreada por líneas horizontales, e Y' por verticales. Entonces, el área cruzada es precisamente la intersección $X' \cap Y'$; la figura muestra que esta área es el complemento de la reunión $X \cup Y$, confirmándose así la segunda ley de dualidad. Este es un argumento convincente para nuestro sentido común, pero no permitido lógicamente, ya que tales demostraciones inductivas están excluidas del razonamiento matemático.

En segundo lugar, podemos considerar separadamente cada uno de los casos posibles para un elemento de I: primero, un elemento en X y en Y; segundo, un elemento en X pero no en Y; y así sucesivamente. Por ejemplo, un elemento del primer tipo está en $X \cap Y$, y por lo tanto no en $(X \cap Y)'$ y tampoco en $X' \cup Y'$, mientras que un elemento del segundo tipo está en $(X \cap Y)'$ y también en Y', y por lo tanto en $X' \cup Y'$. Examinando los otros dos casos posibles, se ve que $(X \cap Y)'$ y $X' \cup Y'$ tienen los mismos elemen-

tos, como afirma la primera ley de dualidad. Hay que observar que para dos clases X e Y, los cuatro casos posibles para un elemento están representados por puntos de las cuatro áreas que muestra el diagrama de Venn de fig. 2, mientras que para tres clases hay ocho casos y ocho áreas (fig. 1). En lógica, a esta imagen de los distintos casos se le llama método del «cuadro de la verdad».

En tercer lugar, podemos usar las mismas definiciones de «cup» y «cap» en la formulación de las leyes. Consideremos la ley distributiva. En ella,

«*b* en $X \frown (Y \smile Z)$» significa «*b* está en X y (también) en Y o en Z»;
«*b* en $(X \frown Y) \smile (X \frown Z)$» significa «*b* está en (ambos) X e Y o en (ambos) X y Z».

Reflexionando un poco se observa que las dos afirmaciones sobre *b* son equivalentes, de acuerdo con el uso ordinario de las conjunciones «y» (simultáneamente) y «o» (alternativa). Esta verificación de la ley distributiva muestra cómo las leyes del álgebra de clases son paráfrasis de las propiedades de las palabras «y», «o» y «no». Si se asumen estas propiedades como básicas, tal como es normal en el razonamiento matemático, se puede probar con ellas el conjunto de leyes que ha quedado expuesto.

Finalmente, el procedimiento algebraico más satisfactorio es elegir como principios básicos algunas de las propiedades que hemos sumariado y, partiendo de éstas, deducir las otras rigurosamente. En otras palabras, puede demostrarse que el conjunto de tales leyes es excesivo, en el sentido de que algunas son consecuencia de las otras. En la elección de las fundamentales se debe atender a la relativa evidencia de las leyes. Esto será desarrollado con algún detalle en los §§ 7-9. Pero por ahora aceptaremos, sin más comentario, el conjunto de todas las leyes señaladas.

EJERCICIOS

1. Verificar la ley distributiva mediante el diagrama de Venn.
2. Emplear el método de subdivisión en casos posibles para verificar las leyes asociativa y conmutativa y el principio de conformidad.
3. Formular de nuevo las leyes de complementación, dualidad e involución mediante los términos «y», «o» y «no», según el tercer método antes descrito.

4. a) Si se trata de una expresión algebraica con cuatro clases, considerando todos los casos que son posibles para un elemento, ¿cuántas posibilidades aparecen?

*b) Dibujar un diagrama para cuatro clases en el cual a cada una de estas posibilidades corresponda una región. (*Sugerencia:* No usar círculos en el diagrama.)

c) Expresar cada región de este diagrama (como en Ejerc. 1, § 1).

5. Probar que las propiedades de intersección y reunión para los conjuntos 0 e I pueden ser deducidas de las de acotación universal y del principio de conformidad.

3. Consecuencias

Variados principios del álgebra de conjuntos pueden deducirse de las leyes ya formuladas. En esta sección señalaremos algunos de los más importantes e inmediatos, consecuencia directa de estas leyes.

Primeramente se verá que hay cuatro modos posibles de relacionar por inclusión dos conjuntos X e Y. Puede tenerse $X \leqslant Y$ e $Y \leqslant X$, en tal caso, por antisimetría, $X = Y$; puede ser $X \leqslant Y$, pero no $Y \leqslant X$, en cuyo caso escribiremos $X < Y$; podemos tener $Y \leqslant X$, pero no $X \leqslant Y$, en cuyo caso escribiremos $X > Y$; y finalmente, cuando no sea $X \leqslant Y$ ni $Y \leqslant X$, diremos que X e Y son incomparables entre sí. La existencia de elementos incomparables es lo que principalmente distingue la relación de inclusión de la de desigualdad entre los números reales.

Las consecuencias de las leyes asociativa y conmutativa han sido ya consideradas en el Cap. I, § 2. La ley asociativa significa esencialmente que podemos formar múltiples intersecciones y reuniones sin usar paréntesis; la ley conmutativa, que podemos permutar términos en cualquier forma, en una expresión que contenga solamente «cups» o solamente «caps».

El efecto de la ley de idempotencia juntamente con las anteriores es, sencillamente, permitir la reducción de grupos en los que un término aparece repetido. En resumen, obtenemos:

LEMA 1. *Por efecto combinado de las leyes asociativa, conmutativa e idempotente, resulta el principio de que la intersección (o la reunión) de un número finito de conjuntos (iguales o distintos) depende sólo de la clase de los conjuntos que intervienen,*

pero no del orden en que se les relacione ni del número de veces que un conjunto esté repetido.

Nuevamente como en el Cap. I, § 5, partiendo de las leyes conmutativa, asociativa y distributiva, podemos obtener leyes distributivas generalizadas, como, por ejemplo,

$$X \frown (Y_1 \smile \ldots \smile Y_n) = (X \frown Y_1) \smile \ldots \smile (X \frown Y_n),$$

$$X \smile (Y_1 \frown \ldots \frown Y_n) = (X \smile Y_1) \frown \ldots \frown (X \smile Y_n),$$

$$(X_1 \smile \ldots \smile X_m) \frown (Y_1 \smile \ldots \smile Y_n) =$$
$$= (X_1 \frown Y_1) \smile (X_1 \frown Y_2) \smile \ldots \smile (X_m \frown Y_n).$$

Más novedad ofrece el

LEMA 2. *La unión y la intersección de clases satisface la ley de absorción:* $X \frown (X \smile Y) = X \smile (X \frown Y) = X$.

Demostración. Evidentemente, $X \smile (X \smile Y) = (X \smile X) \smile Y = X \smile Y$; luego, por el principio de conformidad, $X \frown (X \smile Y) = X$. Para probar que $X \smile (X \frown Y) = X$ se procede igual, intercambiando los «cups» y «caps».

LEMA 3. *Para hallar el complemento de una expresión formada de letras con y sin primas, ligadas por «cups» y «caps» (suponiendo que no haya ningún paréntesis con prima), basta cambiar entre sí los «cups» y los «caps», suprimir la prima en las letras que la tengan y poner prima en las que carecen de ella.*

Así, el conjunto complementario de $(X' \frown Y) \smile (Z \frown W')$ es, según esta regla, $(X \smile Y') \frown (Z' \smile W)$.

Demostración. Si la expresión dada, F, tiene n letras contando las repetidas, el lema es verdadero para $n=1$, ya que $(X)' = X'$ y $(X')' = X$. Por otra parte, no habiendo paréntesis con prima podemos escribir la expresión dada en la forma $F = A \frown B$ o $F = A \smile B$, obteniendo respectivamente $F' = A' \smile B'$ o $F' = A' \frown B'$. Pero las expresiones de A y de B contienen cada una menos letras que la F. Luego, por inducción relativa a n, podemos asegurar que el lema es verdadero para ellas. Y sustituyendo en las expresiones $F' = A' \smile B'$ o $F' = A' \frown B'$, se obtiene el complementario de F según la regla dada.

LEMA 4. *Si* $A \frown X = A \frown Y$ *y* $A \smile X = A \smile Y$, *entonces* $X = Y$.

Demostración. De acuerdo con tales igualdades, y por las leyes de absorción y distributiva, resulta:

$$X = X \frown (X \smile A) = X \frown (Y \smile A) = (X \frown Y) \smile (X \frown A) =$$
$$= (X \frown Y) \smile (Y \frown A) = Y \frown (A \smile X) = Y \frown (A \smile Y) = Y.$$

Como corolario, resulta que el conjunto complementario de otro es único. Pues dado A, será A' el único conjunto que cumple las igualdades $A \frown X = 0$ y $A \smile X = I$.

EJERCICIOS

1. Hallar la expresión complementaria de cada una de las siguientes:
   a) $X \smile Y \smile Z'$;
   b) $(X \smile Y' \smile Z') \frown (X \smile (Y \smile Z'))$;
   c) $X \smile (Y \frown (Z \smile W'))$;
   d) $(X' \smile Y)' \frown (X \smile Y')$.
2. Llevar a cabo la prueba detallada del Lema 3 en el caso especial de la expresión $(X' \frown Y \frown Z') \smile (X \frown Y')$.
3. Simplificar las siguientes expresiones booleanas:
   a) $(X' \frown Y')'$;
   b) $(A \smile B) \smile (C \smile A) \smile (B \smile C)$;
   c) $(X \frown Y) \smile (Z \frown X) \smile (X' \smile Y')'$.
4. Demostrar que $(X \frown Y) \smile (X \frown Y') \smile (X' \frown Y) \smile (X' \frown Y') = I$. ¿Qué significa esto sobre el diagrama de dos círculos?
5. Probar que $X = Y$ si, y sólo si, $(X \frown Y') \smile (X' \frown Y) = 0$.
6. Demuéstrese la siguiente ley de Poretsky: Dados X y T, es $X = 0$ si, y sólo si, $T = (X \frown T') \smile (X' \frown T)$.
7. Probar que: a) $Y \leqslant X'$ si, y sólo si, $X \frown Y = 0$;
   b) $Y \geqslant X'$ si, y sólo si, $X \smile Y = I$.

4. Aplicación a la lógica

El álgebra de clases está íntimamente unida a las propiedades elementales de los juicios o aserciones lógicas.

Los juicios declaratorios (llamados también proposiciones) serán ordinariamente designados por las letras «p», «q» o «r». Estos juicios pueden ser combinados mediante las palabras «y», «o» y «no», produciéndose así nuevas proposiciones, lo mismo que la combinación de clases mediante intersecciones, reuniones y complementos da nuevas clases o conjuntos. Designaremos «o» por «cup», «y» por

«cap» y «no» por prima. Así, si «p» denota al juicio específico «---» y «q» significa otro juicio «···»,

$$«p \smile q» \text{ designa el juicio } «--- \text{ o } \cdots»,$$
$$«p \frown q» \text{ designa el juicio } «--- \text{ y } \cdots»,$$
$$«p'» \text{ designa el juicio } «\text{no } ---».$$

Así, siguiendo, diremos que «$p=q$» significa que las aserciones designadas por las letras «p» y «q» son lógicamente equivalentes. Como es claro, $p \smile q = q \smile p$, de modo que la conjunción «o» es conmutativa. Del mismo modo, el juicio «no (--- y ···)» significa lo mismo que el juicio «o no --- o no ···» ; por lo tanto, $(p \frown q)' = p' \smile q'$. Esta es la primera ley de dualidad del § 2. Del mismo modo se verifica que las conjunciones de los juicios ofrecen las mismas propiedades básicas señaladas en el § 2 para el álgebra de clases. Es conveniente disponer de una locución para designar en general un sistema algebraico de esta naturaleza.

Definición. *Un álgebra booleana (o de Bool) es un conjunto de elementos con las siguientes propiedades:*

I) B *posee dos operaciones binarias, «cup» y «cap», las cuales satisfacen las leyes idempotente, conmutativa, asociativa y distributiva del* § *2.*

II) B *posee una relación denotada por «$\leqslant$», la cual es reflexiva, antisimétrica y transitiva, y satisface el principio de conformidad.*

III) B *contiene dos elementos, 0 e* I, *los cuales son cotas universales y satisfacen las leyes de intersección y unión expresadas en* § *2.*

IV) B *posee la operación unitaria de complementar, la cual obedece las leyes de complementación, dualidad e involución.*

En particular, el álgebra de los subconjuntos de un conjunto I es un álgebra booleana. También lo es el álgebra de los juicios o proposiciones si tomamos los juicios verdaderos como I y los falsos como 0.

Teorema 1. *El álgebra de las proposiciones ligadas por las palabras «y», «o» y «no» es un álgebra booleana abstracta.*

Otra importante conjunción entre juicios está dada por el condicional «si ---, entonces ···». Lo mismo se expresa a veces di-

[illegible] implica ···». Esta conexión entre los dos juicios se [illegible] una flecha, de modo que si «p» designa «---» y «q» [illegible], tendremos:

$$«p \to q» \text{ significa el juicio «si ---, entonces ···».}$$

[illegible] definirse mediante otras conjunciones, ya que «$p \to q$» [illegible] q puede sostenerse hasta que p falle, así que

[illegible] $$(p \to q) = (p' \smile q).$$

[illegible] como $p \smile p' = I$ significa que $p \smile p'$ es equivalente [illegible] una aserción verdadera I. En otras palabras, el juicio [illegible] es verdadero sin importar lo que el juicio componente p [illegible] sido, de modo que «$p \smile p'$» es verdad en virtud de su [illegible] lógica. Esta clase de expresiones son llamadas *tautologías*. Otra importante tautología es

[illegible] $$[(p \to q) \frown (q \to r)] \to (p \to r)$$

[illegible] esencialmente que la implicación es transitiva [illegible] r dan $p \to r$). Para demostrar que (2) es una tauto[illegible] sustituir cada flecha tal como indica la definición (1). [illegible] (2) resulta

$$[(p' \smile q) \frown (q' \smile r)]' \smile (p' \smile r).$$

[illegible] del término entre corchetes puede hallarse según [illegible], obteniéndose

$$[illegible] \smile r') \smile p' \smile r = [p' \smile (p \frown q')] \smile [r \smile (q \frown r')] =$$
$$= [(p' \smile p) \frown (p' \smile q')] \smile [(r \smile q) \frown (r \smile r')] =$$
$$= p' \smile q' \smile r \smile q = p' \smile r \smile (q' \smile q) = I.$$

[illegible], (2) es una tautología. Otras tautologías impor[illegible]

[illegible] $$(p \smile p) \to p, \quad p \to (p \smile q), \quad (p \smile q) \to (q \smile p),$$
$$(p \to q) \to [(r \smile p) \to (r \smile q)].$$

[illegible] tautologías son importantes porque es posible de[illegible] que cualquier otra puede derivarse de ellas mediante los [illegible] de demostración siguientes:

Sustitución: En una tautología, cualquier letra puede ser reemplazada por una combinación de letras fija.

Inferencia (o ilación): Si X y $X \to Y$ son tautologías, también lo es Y.

El posterior desarrollo de la lógica matemática requiere un cuidadoso análisis de las reglas demostrativas de este tipo. Además de las conexiones entre juicios representadas por las operaciones del álgebra de Bool, deben tenerse en cuenta frases como estas: «existe una x tal, que ...» y «para todos los x, ...», en que se envuelven los cuantitativos «existe» y «para todos».

EJERCICIOS

1. Demostrar que los cuatro juicios de (3) son tautologías.
2. Probar que son tautologías los siguientes juicios:
   a) $p' \to (p \to q)$;
   b) $[p \frown (p \to q)] \to q$;
   c) $p \to [q \to (p \to q)]$.
3. ¿Cuáles de las siguientes expresiones son tautologías?
   a) $p \to p'$;
   b) $(p \to q) \to (q' \to p')$;
   c) $(p \to q') \to (q' \to p)$.
4. Si $f(p, q, \ldots)=0$, para todos los $p, q, \ldots,$ el juicio $f(p, q, \ldots)$ es llamado un *absurdo*. Probar que cada uno de los siguientes juicios es un absurdo:
   a) $p \frown p'$;
   b) $q' \frown p \frown (p \to q)$;
   c) $(p' \to p) \frown (p \to p')$;
   d) ¿Es absurdo $p \to p'$?
5. Definimos «p si, y sólo si, q» por $p \leftrightarrow q = [(p \to q) \frown (q \to p)]$. Establecer y demostrar dos tautologías implicando el $\leftrightarrow$.
6. Probar que toda álgebra booleana tiene una subálgebra que consta de dos elementos.
7. Si I es un conjunto que contiene dos elementos nada más, escribir la «tabla de multiplicar» para las operaciones de intersección y reunión en el álgebra de todas las subclases de I.
8. Definir la noción de isomorfismo entre dos álgebras booleanas. Si dos clases finitas I e I' tienen el mismo número de elementos, probar que el álgebra de todas las subclases de I es isomorfa con el álgebra de todas las subclases de I'.
9. Probar que para cualquier entero positivo n existe un álgebra booleana con 2^n elementos.
10. Probar que el álgebra booleana de todos los subconjuntos de una clase de n elementos tiene exactamente $n!$ automorfismos.

5. Forma canónica de las funciones booleanas

Consideremos cualquier álgebra booleana como un álgebra de elementos $a, b, c, \ldots$ Si B es un álgebra de clases, lo dicho significa que estas clases serán consideradas ahora simplemente como elementos sobre los que pueden actuar las operaciones «cup» y «cap»: llamamos $a \cap b$ la intersección de a y b, $a \cup b$ la reunión de a y b.

En las secciones precedentes hemos considerado repetidamente funciones construidas con las operaciones elementales «cap», «cup» y «prima». Será conveniente que en lo sucesivo las llamemos «funciones booleanas» o «polinomios booleanos». Tienen cierta analogía evidente con las funciones polinomiales ordinarias, advirtiendo, desde luego, como diferencia, que la teoría de los polinomios booleanos en una variable es trivial: hay solamente los cuatro polinomios x, x', $0 = x \cap x'$, $I = x \cup x'$.

Un polinomio booleano con n variables puede llevarse a una forma canónica aplicando sistemáticamente las leyes del álgebra booleana. El desarrollo de la explicación será ilustrado con el ejemplo del polinomio $F = [(x \cap y)' \cup z] \cap (z \cup x)'$.

Ante todo, si una prima se encuentra fuera de un paréntesis puede llevarse al interior por una aplicación de la ley de dualidad (ver Lema 8 de § 3). Hecho esto con todas las primas que afecten a paréntesis, el polinomio se transforma en una expresión envolviendo letras con y sin prima, unidas por «caps» y «cups». Por ejemplo,

$$F = [(x \cap y)' \cup z] \cap (z \cup x)' = [x' \cup y' \cup z] \cap (z' \cap x').$$

En segundo lugar, donde haya un «cap» fuera de un paréntesis que encierra un «cup», puede aquel signo introducirse dentro, aplicando la ley distributiva como en $c \cap (a \cup b) = (c \cap a) \cup (c \cap b)$. Así resulta un polinomio, en el que todos los «caps» están aplicados antes que los «cups», esto es, la expresión es una reunión de términos en que cada término T es una intersección de letras con prima y sin prima. En el ejemplo anterior,

$$F = [x' \cup y' \cup z] \cap (z' \cap x') =$$

$$= (x' \cap z' \cap x') \cup (y' \cap z' \cap x') \cup (z \cap z' \cap x').$$

En tercer lugar, atenderemos a que ciertas expresiones pueden simplificarse. Si una letra x aparece dos veces en un término, la repetición puede ser omitida, por ser $x \cap x = x$. Si una letra aparece en el mismo término con prima y sin ella, como en el primer término del polinomio anterior, todo el término es cero, ya que $x \cap a \cap x' = 0$ para cualquier a. Finalmente, supongamos que en uno de los términos, P, no figura alguna de las letras dadas ni con prima ni sin ella; pero si falta la x, por ejemplo, añadiendo el factor $x \cup x' = I$ obtendremos $P = P \cap (x \cup x') = (P \cap x) \cup (P \cap x')$. Con esto, el término P puede ser reemplazado por dos nuevos términos conteniendo cada uno la letra que faltaba. En el resultado, cada término contiene a cada letra una vez, y solamente una vez. En nuestro ejemplo,

$$F = (y \cap z' \cap x) \cup (z' \cap x) =$$
$$= (y \cap z' \cap x) \cup (y \cap z' \cap x) \cup (y' \cap z' \cap x).$$

Omitiendo la repetición del primer término queda finalmente $F = (y \cap z' \cap x) \cup (y' \cap z' \cap x)$.

Por este procedimiento, cada polinomio booleano en tres variables, x, y, z, puede ser reducido a cero o a una reunión de varios de los términos siguientes:

$$\begin{array}{llll} x \cap y \cap z, & x' \cap y \cap z, & x \cap y' \cap z, & x \cap y \cap z', \\ x' \cap y' \cap z, & x' \cap y \cap z', & x \cap y' \cap z', & x' \cap y' \cap z'. \end{array} \tag{4}$$

No es accidental el hecho de que estos ocho polinomios representen los ocho espacios en que los tres círculos de la figura 1 dividen al cuadrado. Esto significa geométricamente, que cada combinación booleana de los tres círculos X, Y y Z será la unión de alguna selección de los ocho espacios del diagrama.

Los términos obtenidos al fin de la reducción, tal como los indicados en (4), serán llamados polinomios booleanos mínimos. En otras palabras, un polinomio *mínimo* en n variables $x_1, x_2, \ldots, x_n$ es una intersección de n letras, en la que la letra de lugar i es x_i o x_i'. Hay, pues, los 2^n polinomios mínimos siguientes:

$$(x_1 \cap x_2 \cap \ldots \cap x_n), (x_1' \cap x_2 \cap \ldots \cap x_n), \ldots, (x_1' \cap x_2' \cap \ldots \cap x_n').$$

TEOREMA 2. *Existe una, y sólo una, manera de representar un polinomio booleano dado como cero o como reunión de polinomios mínimos.*

(Desde el momento en que 0 puede ser considerado como la reunión del conjunto vacío de polinomios mínimos, la excepción mencionada es sólo aparente.)

El hecho de que por lo menos existe una tal representación ha sido probado por las consideraciones que acabamos de desarrollar. Para demostrar la unicidad, en otras palabras, para probar que las uniones de distintos conjuntos de polinomios mínimos representan distintas funciones booleanas, construiremos un sencillo modelo. Sea I el conjunto de los decimales con n cifras en el sistema diádico o, lo que es igual, el conjunto de ordenaciones de n términos iguales a 0 o a 1. Llamemos X_i al conjunto de todas estas ordenaciones cuya cifra de lugar i sea la unidad. Cada polinomio mínimo P_k ($k=1, 2, 3, \ldots, 2^n$) representa un elemento de la clase I, precisamente aquel cuya cifra i-ésima es 1 o 0, según que el i-ésimo término del polinomio sea x_i o x_i'. Consecuentemente, la reunión de distintos conjuntos de polinomios mínimos representa diferentes conjuntos de tales decimales y distintos «polinomios booleanos canónicos» representan funciones distintas. Esto completa la demostración del Teorema 2.

COROLARIO. *Existen precisamente 2^{2^n} funciones booleanas distintas con* n *variables.*

Se puede ahora sustituir la manipulación al azar de los polinomios booleanos por un procedimiento sistemático. La verdad o falsedad de una igualdad dada $E_1=E_2$ en álgebra booleana, puede ser establecida definitivamente por simple reducción de cada miembro a su forma canónica.

EJERCICIOS

1. Reducir cada una de las siguientes expresiones a su forma canónica:
   a) $(x \smile y) \frown (x' \frown y)'$;
   b) $(x \smile y) \frown (y \smile z) \frown (x \smile z)$.
2. Comprobar cada una de las siguientes igualdades por reducción de ambos miembros de su forma canónica:
   a) $[x \frown (y \smile z)']' = (x \frown y)' \smile (x \frown z)$;
   b) $x = (x' \smile y')' \smile [z \smile (x \smile y)']$.
3. Demostrar que todo polinomio booleano admite otra forma canónica que consiste en la intersección de varios polinomios con prima. Describir cuidadosamente estos polinomios y probar que son los complementos de los

polinomios mínimos. ¿Subsiste un teorema análogo al que afirma que un polinomio ordinario sobre un campo tiene una expresión única como producto de polinomios irreducibles?

4. Usar la forma canónica del Ejerc. 3 para la comprobación del Ejerc. 2 a).
5. Probar que la forma canónica de $f(x, y)$ es

$$f(x, y)=[f(I, I) \frown x \frown y] \smile [f(I, 0) \frown x \frown y'] \smile [f(0, I) \frown x' \frown y] \smile [f(0, 0) \frown x' \frown y'].$$

6. Demostrar que la intersección de dos polinomios mínimos distintos es 0.
7. Desarrollando $I=(x_1 \smile x_1') \frown \ldots \frown (x_n \smile x_n')$ por la ley distributiva generalizada, demostrar que I es la reunión de todos los polinomios mínimos.
8. Utilizando el Ejerc. 7 y la igualdad $x_i=x_i \frown I$ probar que cada x_i es igual a la reunión de todos los polinomios mínimos cuyo término i-ésimo es x_i.
9. a) Denotaremos por $\bigvee_S P_i$ la reunión de todos los polinomios mínimos en el conjunto S. Probar que

$$\bigvee_S P_i \smile \bigvee_T P_i = \bigvee_{S \smile T} P_i, \qquad \bigvee_S P_i \frown \bigvee_T P_i = \bigvee_{S \frown T} P_i.$$

b) Probar que las anteriores fórmulas son ciertas cuando se define como 0 el conjunto $\bigvee_0 P_i$, esto es, el conjunto vacío de polinomios mínimos.

*10. Utilizando los Ejercicios 7 y 9, probar que $\left(\bigvee_S P_i\right)' = \bigvee_{S'} P_i$. (*Sugerencia:* Repasar el final de § 3.)
11. Utilizando solamente los resultados de los Ejercicios 8, 9 y 10, dar una nueva prueba de que cada polinomio booleano puede ser expresado como una reunión de polinomios mínimos.

6. Aplicaciones del álgebra booleana

Los métodos del álgebra de Bool pueden ser utilizados para formalizar los procesos del razonamiento elemental. Supongamos que nos han dado las premisas «Algunas leyes son complicadas», «Ninguna ley confusa es satisfactoria» y «Todas las leyes complicadas son confusas». Designemos con I la clase de todas las leyes, con B la clase de todas las leyes complicadas, con C la clase de las leyes confusas y con D la clase de las leyes satisfactorias. Con esta notación, las premisas se leen :

$$B > 0, \quad C \frown D = 0, \quad B \leqslant C.$$

Puesto que $C \frown D = 0$, $C = C \frown (D \smile D') = (C \frown D) \smile (C \frown D')$ es $0 \smile (C \frown D') = C \frown D'$, y por lo tanto, $C \leqslant D'$. (¿Qué leyes del álgebra booleana se han utilizado en lo anterior?)

Combinando esto con nuestra hipótesis, tendremos:

$$0 < B \leqslant C \leqslant D', \quad \text{y por tanto,} \quad D' > 0,$$

que se enuncia: «Algunas leyes no son satisfactorias».

Por tales procedimientos, todos los silogismos de la lógica de Aristóteles pueden ser formulados en términos de álgebra booleana.

Esta álgebra se presta también al manejo de condiciones de complicado enunciado, tales como las que aparecen en algunos tipos de pólizas de seguros. Otra aplicación se refiere al estudio de las clasificaciones numéricas.

Problema (Exámenes para actuarios, 1935, Parte 5.ª, cuestión 9B). Ciertos datos obtenidos del estudio de un grupo de 1000 empleados en una fábrica de hilaturas, atendiendo a su raza, sexo y estado civil, son como siguen: 525 individuos de color, 312 hombres, 470 casados, 42 hombres de color, 147 individuos casados y de color, 86 hombres casados y 25 hombres de color casados. Comprobar la clasificación para determinar si los números dados en los distintos grupos son ciertos.

Los datos dan los números de elementos que hay en ciertas clases y en sus intersecciones. Sean

C = individuos de color,
M = individuos de sexo masculino,
W = individuos casados (y casadas).

Denotemos con $N(A)$ el número de elementos en una clase A. Por la definición de reunión y de intersección, este operador N tiene la propiedad de que

$$N(A \smile B) = N(A) + N(B) - N(A \frown B).$$

Aplicando dos veces esta ley, podemos calcular:

$$N(A \smile B \smile D) = N(A) + N(B \smile D) - N[(A \frown B) \smile (A \frown D)] =$$
$$= N(A) + N(B) + N(D) - N(B \frown D) - N(A \frown B) - N(A \frown D) +$$
$$+ N(A \frown B \frown D).$$

La fórmula que ha resultado es simétrica en A, B y D. Reemplazando estas letras por C, M y W, respectivamente, y atendiendo a los datos, se obtiene:

$$N(C \smile M \smile W) = 525 + 312 + 470 - 42 - 147 - 86 + 25 = 1057,$$

resultado superior al número total de empleados.

EJERCICIOS

1. En una encarnizada batalla, por lo menos el 70 por ciento de los combatientes pierde un ojo; al menos un 75 por ciento, una oreja; al menos un 80 por ciento, un brazo, y al menos un 85 por ciento, una pierna. ¿Cuántos por lo menos han perdido las cuatro cosas? (Lewis Carroll.)
2. ¿Puede creerse a un investigador que informa que, de cada 1000 habitantes, a 816 les gusta el azúcar; a 723, el helado; a 645, los pasteles, y asimismo que a 562 les gusta el azúcar y el helado; a 463, el azúcar y los pasteles, y a 470, los pasteles y el helado, y sólo a 310 les gustan las tres cosas?
3. Tres monedas, una de cobre, otra de níquel y otra de plata, son lanzadas simultáneamente al aire, y esto 100 veces. La de cobre muestra cara en 70 ocasiones, la de níquel muestra cara en 50 pruebas y la de plata muestra cara en 56 pruebas. La de cobre y la de níquel han mostrado simultáneamente cara 31 veces y la de níquel y de plata han presentado tal circunstancia 28 veces. Probar que las tres monedas han mostrado cara simultáneamente por lo menos 9 veces y que las tres han mostrado simultáneamente cruz a lo más 11 veces.

'. Ordenaciones parciales

En las secciones siguientes analizaremos más atentamente las eyes de álgebra de clases, que enunciamos en § 2. Especialmente lemostraremos que ciertas propiedades de la reunión e intersección pueden deducirse de las de la relación de inclusión, y que estas propiedades nos llevan a otros sistemas que no son un álgebra booleana. De este modo entraremos en la materia propia de la teoría de redes lattice theory). Y de paso tendremos una justificación mejor para os enunciados del § 2 y se simplificará el concepto de álgebra booeana.

Las leyes más fundamentales del álgebra de Bool son las refleiva, antisimétrica y transitiva, relativas a la inclusión. Ellas son videntemente válidas para los subconjuntos de cualquier sistema, que se caracterizan por ciertas propiedades dadas. Así se cumplen, or ejemplo, para los subgrupos de cada grupo (o para los subgruos normales), los subgrupos de cada campo, los subespacios de ada espacio lineal y así sucesivamente, aun cuando estos subsisteoas no constituyan álgebras booleanas. También son válidas dichas eyes para las relaciones «menor o igual que», $x \leqslant y$, entre números reales y para la relación de divisibilidad, $x \mid y$, entre números aturales, etc.

Tales ejemplos sugieren el concepto abstracto de sistema ordenado parcialmente. Esto significa un sistema cuyos elementos están relacionados por una ley reflexiva, antisimétrica y transitiva. Para cualquier sistema con una relación de este tipo, $a \leq b$ (leer «b incluye a a»), podemos poner, por definición, $a < b$, para significar que $a \leq b$ y además $a \neq b$. Diremos que b es una *cobertura* de a, si es $a < b$ y además, para ningún x puede ser $a < x < b$.

Los sistemas ordenados parcialmente con un número finito de elementos pueden ser convenientemente representados por diagramas. Cada elemento del sistema está representado por un pequeño círculo de modo que el círculo para a quede encima del círculo

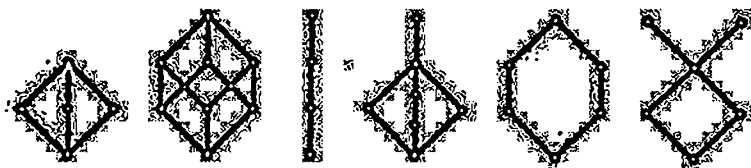


Figura 3

para b si $a > b$. Además, cuando a sea una cobertura de b se traza desde a un segmento que descienda hasta b. Se puede reconstruir la relación $a \geq b$ a la vista del diagrama; será $a \geq b$ si es posible pasar desde b hasta a siguiendo una línea ascendente de segmentos del diagrama, y sólo en tal caso.

Por ejemplo, en la figura 3 el primer diagrama representa el sistema de todos los subgrupos del grupo del rectángulo; el segundo, el álgebra booleana de todos los subconjuntos de un conjunto de tres puntos; el tercero, los números 1, 2, 4 y 8 bajo la relación de divisibilidad. Los otros han sido trazados a capricho y demuestran que es posible construir sistemas abstractos parcialmente ordenados con el simple trazado de diagramas. En el Cap. VI, § 8, se encuentra el diagrama de todos los subgrupos del grupo del cuadrado.

Es claro que en un sistema ordenado parcialmente, la relación $\geq$ es también reflexiva, antisimétrica y transitiva (lo cual resulta sin más que leer de derecha a izquierda en el enunciado de estas leyes). Por consiguiente, toda proposición que pueda ser probada a partir de los postulados que definen un sistema parcialmente ordenado respecto a la relación $a \leq b$ puede ser establecida exactamente con los mismos razonamientos si siempre la relación $a \leq b$

viene reemplazada por la relación $a \geq b$, y viceversa. Esto constituye el

PRINCIPIO DE DUALIDAD. *Cualquier teorema cierto en todo sistema ordenado parcialmente sigue siendo cierto si los símbolos $\leq$ y $\geq$ se intercambian mutuamente en el enunciado.*

Hay que recalcar que este principio no es un teorema relativo a los sistemas ordenados parcialmente, en el sentido ordinario de la palabra, antes bien es un teorema relativo a los teoremas. Como tal, pertenece al estudio de la «metamatemática», estudio que ha llegado a ser muy importante en la moderna lógica simbólica.

EJERCICIOS

1. Mostrar con detalle que el segundo diagrama de la fig. 3 representa el álgebra de todos los subconjuntos de un conjunto de tres puntos.
2. Dibujar un diagrama para cada uno de los siguientes sistemas parcialmente ordenados:
   a) El álgebra booleana de todos los subconjuntos de un conjunto de cuatro puntos;
   b) El sistema de todos los subgrupos de un grupo cíclico de orden 12;
   c) El sistema de todos los subgrupos del grupo cuaternio;
   d) Los números 1, 2, 3, 4, 6, 8, 12, 24, bajo relación de divisibilidad;
   e) El sistema de todos los subgrupos de un grupo cíclico de orden 54;
   f) El sistema de todos los subanillos del anillo J_n de los enteros módulo 10.
3. Discutir la interrelación de los sistemas ordenados parcialmente de las partes b), e), f) del Ejercicio 2.
4. ¿Cuáles de los siguientes conjuntos son sistemas ordenados parcialmente?
   a) Todos los subcampos del campo R^* de los números reales, bajo la relación de inclusión;
   b) Todos los pares de números (a, b) si $(a, b) \leq (a', b')$ significa que $a \leq a'$ y $b \leq b'$;
   c) Todos los pares de números cuando $(a, b) \leq (a', b')$ significa que $a < a'$ o bien, que si $a = a'$ es $b \leq b'$;
   d) Todos los pares de números si $(a, b) \leq (a', b')$ significa que $a + b \leq a' + b'$ y $b \leq b'$;
   e) Todos los subdominios de un dominio de integridad dado, bajo la relación de inclusión;
   f) Todos los polinomios en $F[x]$, si $f(x) \leq g(x)$ significa que $f(x)$ divide a $g(x)$.
5. Consideremos un sistema de elementos con la relación $a < b$, la cual es transitiva e irreflexiva ($a < a$ no es cierto jamás). Si $a \leq b$ significa que $a < b$ o $a = b$, probar que se engendra un sistema ordenado parcialmente.

8. Redes (Lattices) (*)

El principio de conformidad muestra cómo se define la inclusión refiriéndola a la reunión o intersección; ahora demostraremos que, recíprocamente, pueden definirse la reunión e intersección mediante la inclusión. Precisamente, $x \smile y$ es la menor entidad que contiené a ambas x e y, mientras que $x \frown y$ es la mayor entidad contenida en ambas x e y. Esta observación se debe a C. S. Peirce; la estableceremos con más precisión como sigue.

Se entiende por «cota inferior» de un conjunto parcialmente ordenado X a un elemento a satisfaciendo a la condición $a \leqslant x$ para todo $x \epsilon X$. Llamaremos «cota inferior máxima» (c. i. m.) a una cota inferior que contiene a cualquier otra cota inferior: o sea, una cota inferior c tal, que $c \geqslant a$ para toda otra cota inferior a. Claramente veremos que las cotas inferiores máximas, si existen, son únicas; sean a y b dos c. i. m. del mismo conjunto X. Entonces debe ser $a \geqslant b$ y $b \geqslant a$ y por consiguiente $a = b$.

Dualmente se definen el concepto de «cota superior» y el de «cota superior mínima» (c. s. m.) y se prueba su unicidad en el caso de que existan. ¡Estamos aquí aplicando el principio metamatemático de dualidad! Por lo tanto, es legítimo hablar de *la* c. i. m. y de *la* c. s. m. de un conjunto de elementos, siempre que estas cotas existan.

LEMA 1. *En un álgebra booleana, la intersección* $x \frown y$ *y la reunión* $x \smile y$ *son la c. i. m. y la c. s. m. respectivamente del conjunto que forman los dos elementos* x *e* y.

Demostración. Puesto que $x \frown x \frown y = x \frown y$ e $y \frown x \frown y = x \frown y$, el principio de conformidad demuestra que $x \frown y$ es una cota inferior de x e y. Es la máxima cota inferior, ya que $z \leqslant x$ y $z \leqslant y$ implica $z = x \frown z = x \frown (y \frown z)$, y también $z \leqslant x \frown y$, otra vez por el principio de conformidad. La prueba se completa por dualidad.

Resulta como corolario que un álgebra booleana es (entre otras cosas) un sistema ordenado parcialmente en el cual dos elementos

(*) En vez de «red», algunos autores dicen «retículo», y otros, «estructura», y también «huso». En realidad, lo más acertado sería dejar sin traducir el nombre original *(lattice)*, que se ha hecho internacional. (N. del T.)

cualesquiera tienen una c. i. m. y una c. s. m. De un tal sistema se dice que tiene «estructura de red». Ahora demostraremos que hay muchos sistemas ordenados parcialmente que no son álgebras booleanas, pero que tienen estructura de red.

Así, consideremos el sistema $L(G)$ de todos los subgrupos de un grupo G. La intersección $S \cap T$ de dos de tales subgrupos es también un subgrupo y evidentemente el máximo subgrupo contenido en S y en T. Por otra parte, el subgrupo $S \cup T$ «engendrado» por S y T (Cap. VI, § 8) es el menor subgrupo en $L(G)$ que contiene a S y a T.

Análogamente, consideremos el sistema $L(V)$ de todos los subespacios de un espacio lineal V. La intersección y la combinación lineal de dos de tales subespacios S y T son las c. i. m. y la c. s. m. de S y T en $L(V)$.

Asimismo, el mayor y el menor de dos números reales son su c. i. m. y su c. s. m. bajo la relación de desigualdad. Finalmente, el máximo común divisor y el mínimo común múltiplo de dos números naturales son, respectivamente, la c. i. m. y la c. s. m. de ambos bajo la relación de divisibilidad.

Con todos estos ejemplos en la mente, no será difícil adquirir el concepto de «red». ¡Una red es un sistema ordenado parcialmente en el cual dos elementos cualesquiera tienen una c. i. m. y una c. s. m.! En toda red, la c. i. m. de a y b se designa por $a \cap b$ y la c. s. m. por $a \cup b$.

EJERCICIOS

1. ¿Qué diagramas de la fig. 3 representan redes?
2. Dibujar otros dos diagramas de sistemas ordenados parcialmente, los cuales no sean redes.
3. ¿Cuáles de los ejemplos del Ejerc. 4, § 7, representan redes?
4. Demostrar que la red de todos los subgrupos del grupo del rectángulo no es un álgebra booleana.
5. Ilustrar el principio de dualidad, desarrollando detenidamente la prueba de la segunda parte del Lema 1, según la prueba expuesta para la primera.
6. Probar que toda red que tiene sólo un número finito de elementos posee dos elementos 0 e I satisfaciendo $0 \leqslant x \leqslant I$ para todo elemento x.
7. Probar que todo sistema finito ordenado parcialmente con 0 e I es una red siempre que entre elementos a_1, a_2, b_1, b_2, con $a_i \geqslant b_j$ (i, $j = 1, 2$), exista un elemento c tal, que $a_i \geqslant c \geqslant b_j$ para todo i y j.

*8. Probar que todos los subgrupos normales de un grupo dado G forman una red. (*Sugerencia:* Probar primero que la intersección y reunión de dos subgrupos normales son asimismo normales.)

9. Identidades en las redes

Vamos a ver ahora que gran parte del álgebra de clases se aplica a las redes en general.

Teorema 3. *En una red cualquiera, son válidas las leyes idempotente, conmutativa, asociativa y de absorción, así como el principio de conformidad.*

Demostración. Por el principio de dualidad nos bastará probar estas leyes para la c.i.m. La conmutativa es obvia por la simetría de la definición; la ley asociativa resulta de que tanto $x \cap (y \cap z)$ como $(x \cap y) \cap z$ son la c.i.m. de los tres elementos x, y, z. La ley idempotente es trivial, por sustitución en la definición. Para el principio de conformidad, consideremos primero que $x \leq y$; entonces cualquier z con $z \leq x$ y $z \leq y$ satisface $z \leq x$, y como también es $x \leq x$ y $x \leq y$, resulta que x satisface a la definición de c.i.m. $x \cap y$. Inversamente, si $x = x \cap y$, entonces x es una cota inferior de y, así que $x \leq y$; esto prueba el principio de conformidad. La ley de absorción sigue ahora por una simple repetición de razonamiento del Lema 2, §3.

La ley distributiva no ha sido mencionada en el enunciado del Teorema 3 por la simple razón de que no se satisface en todas las redes. Por ejemplo, no es válida cuando x, y, z son los tres subgrupos de orden 2 en el grupo del cuadrilátero (fig. 3, primer diagrama).

Teorema 4. *En cualquier red son válidas las leyes semidistributivas*

$$x \cup (y \cap z) \geq (x \cup y) \cap (x \cup z),$$
$$x \cap (y \cup z) \leq (x \cap y) \cup (x \cap z).$$

Además, cualquier ley distributiva implica su dual.

Demostración. El trabajo de la demostración se reduce a la mitad por el principio de dualidad. Atendiendo a la primera ley semidistributiva, se observa que los dos términos del primer miembro tienen como cota inferior a cualquiera de los dos términos del segundo miembro; por lo tanto, la c.i.m. de x e $y \cap z$ es una cota superior para $x \cup y$ y $x \cup z$, y por lo tanto, para su c.s.m.

$(x \cap y) \cup (x \cap z)$. Por otra parte, partiendo de una ley distributiva, desarrollando obtenemos

$$(x \cup y) \cap (x \cup z) = [(x \cup y) \cap x] \cup [(x \cup y) \cap z] =$$
$$= x \cup (x \cap z) \cup (y \cap z) = x \cup (y \cap z),$$

que es la otra ley distributiva.

Las redes en las que es válida la ley distributiva se llaman «redes distributivas». Se observa sin demostración que los números reales ordenados por su magnitud y los números naturales ordenados bajo la relación de divisibilidad son redes distributivas.

Teorema 5. *En una red distributiva, el complemento es único (cuando existe) y satisface las leyes de involución y dualismo.*

Demostración. La unicidad está probada en el Lema 1, § 3, dependiendo de la distributividad (estudiar la red de los subgrupos del grupo del cuadrilátero, donde los complementos no son únicos). La simetría de la condición de complementación asegura la ley de involución $(x')' = x$. Para deducir la ley de dualismo obsérvese que

$$(a \cup b) \cap (a' \cap b') = (a \cap a' \cap b') \cup (b \cap a' \cap b') = O,$$
$$(a \cup b) \cup (a' \cap b') = (a \cup b \cup a') \cap (a \cup b \cup b') = I,$$

luego $a' \cap b' = (a \cup b)'$. Ahora empléese el principio de dualidad.

Como corolario de los resultados precedentes deducimos que, para probar que un álgebra de clases es booleana, basta establecer que: I) la inclusión de conjuntos es reflexiva, antisimétrica y transitiva; II) la reunión de dos conjuntos es el menor conjunto que los contiene a ambos y dualmente para la intersección; III) $S \cap (T \cup U) = (S \cap T) \cup (S \cap U)$ idénticamente; IV) cada conjunto S tiene un «complemento» S' tal, que $S \cap S' = O$, $S \cup S' = I$. Esto prueba también el

Teorema 6. *Un álgebra booleana es una red distributiva que contiene dos elementos O e I con $O \leq a \leq I$ para todo a, y en la cual cada a tiene un complementario a' satisfaciendo las condiciones $a \cap a' = O$, $a \cup a' = I$.*

Las álgebras booleanas pueden ser caracterizadas por otros muchos sistemas de postulados. Uno de ellos es el indicado en el siguiente ejercicio 6.

EJERCICIOS

1. Enunciar y demostrar el principio de dualidad para álgebras de Bool.
2. Mostrar detalladamente que la prueba de la segunda ley de dualismo en el Teor. 5 es exactamente dual de la prueba allí dada.
3. Mostrar que la red de todos los subgrupos del grupo del cuadrilátero no es distributiva.
4. Llamamos *cadena* a un conjunto simplemente ordenado (es decir, un conjunto ordenado parcialmente en el cual todo x e y verifican o bien $x \geqslant y$ o bien $y \geqslant x$). a) Demostrar que cualquier cadena es una red distributiva b) Probar que una red es una cadena si, y sólo si, todos sus subconjuntos son subredes.
5. Una red se llama *modular* si, y sólo si, $x \geqslant z$ implica siempre $x \frown (y \smile z) = (x \frown y) \smile z$.
   a) Probar que toda red distributiva es modular.
   b) Construir el diagrama de una red sencilla que no sea modular.
   c) Probar que cada una de las siguientes redes es modular: 1) todos los subespacios de un espacio vectorial; 2) todos los subgrupos de un grupo abeliano; 3) todos los subgrupos normales de cualquier grupo.
   d) Demostrar que en una red modular, $x \leqslant z$ implica siempre $x \smile (y \frown z) = (x \smile y) \frown z$. De aquí se infiere que el principio de dualidad se cumple en las redes modulares.
*6. Si L es una red con las cotas universales 0 e I, en la cual cada elemento a tiene un complemento a' con las propiedades

$$x \leqslant a' \text{ si, y sólo si, } a \frown x = 0,$$
$$y \geqslant a' \text{ si, y sólo si, } a \smile y = I,$$

demostrar que L es un álgebra booleana. (*Sugerencia:* Para demostrar la primera ley distributiva basta probar que $e \equiv [a \frown (b \smile c)] \frown [(a \frown b) \smile (a \frown c)]' = 0$. Expresar e como una intersección y considérense particularmente los términos.)
*7. En un álgebra booleana, la diferencia simétrica de dos elementos x e y se define por la expresión $x + y = (x \frown y') \smile (x' \frown y)$.
   a) ¿Qué significa esto cuando x e y son conjuntos? Dibujar una figura.
   b) Demostrar que $x + y$ es asociativa, conmutativa y con elemento nulo.
   c) Imaginando la diferencia simétrica como una suma y la intersección como un producto, probar que toda álgebra booleana es un anillo conmutativo con elemento unidad.
*8. Sea L un sistema de elementos con dos operaciones binarias («cup» y «cap»), las cuales son asociativas y conmutativas y satisfacen a la ley de absorción. Mediante el principio de conformidad, definir una relación de inclusión en L y probar así que L es una red en la cual las operaciones dadas, «cup» y «cap», son c. s. m. y c. i. m. respectivamente.

CAPITULO XII

Aritmética transfinita

1. Números y conjuntos

El presente capítulo se ocupa de la articulación entre los conceptos de número y de conjunto; esto corresponde al método que llamaremos *cardinal* de introducir los números enteros; en cambio, el método llamado *ordinal* atiende fundamentalmente a la posición del número en la conocida sucesión «uno, dos, tres, ...». Nuestro objeto es deducir del método cardinal algunas propiedades de los números infinitos (¡entre ellas, su definición precisa!), los cuales han mostrado ser de importancia capital en las matemáticas modernas.

El método cardinal ha resultado también importante en el estudio crítico de los fundamentos de las matemáticas. Llevado a su extremo lógico, proporciona la definición de los números mediante los conjuntos, y así se reduce el número de términos no definidos que deben utilizarse en matemáticas. Pero este programa implica una reelaboración de las ideas básicas, ajustándolas con una meticulosa atención, a la que no podemos dedicarnos en el presente texto. Por lo tanto, supondremos que al lector le son familiares el concepto de número natural y el de *conjunto* o clase, y partiremos de ellos.

Las relaciones entre números y conjuntos están cimentadas en la siguiente definición:

DEFINICIÓN. *Sea* n *un entero positivo. Se dirá que un conjunto* S *tiene número cardinal* n [*en símbolos*, o(S)=n] *si, y sólo si,*

existe una correspondencia biunívoca entre los elementos de S *y los enteros 1, 2, 3, ...,* n (*).

Esto significa que los elementos de S pueden ser designados por $s_1, s_2, \ldots, s_n$, siendo s_k el elemento de S que corresponde al entero k. Dicho de otro modo, lo anterior expresa que es posible contar los elementos de S contando hasta n y contando cada elemento una sola vez. Como corolario resulta que, si dos clases S y T tienen el mismo número cardinal, habrá una correspondencia biunívoca entre ellas, esto es, la correspondencia $s_1 \leftrightarrow t_1, \ldots, s_n \leftrightarrow t_n$. Pero no es evidente a priori que un conjunto no pueda tener dos números cardinales distintos, es decir, que contándolos en orden distinto no se pueda llegar a diversos números como total de elementos. Vamos a demostrar esta unicidad, estableciendo primero un resultado algo más general.

TEOREMA 1. *Sean* m *y* n *dos enteros positivos. Existirá una correspondencia biunívoca entre el conjunto 1, ...,* m *y un subconjunto propio de la clase 1, ...,* n *si, y sólo si, es* m < n.

Demostración. Si $m < n$, la correspondencia $1 \leftrightarrow 1$, $2 \leftrightarrow 2, \ldots,$ $m \leftrightarrow m$ es de la naturaleza requerida. Esta primera mitad del teorema es obvia, pero la inversa debe analizarse con más cuidado.

La inversa es trivial cuando $m=1$, ya que 1 es el menor entero positivo; por lo tanto, podemos aplicar la inducción sobre m. En efecto, supongamos que existe una correspondencia biunívoca $1 \leftrightarrow f(1), \ldots, m \leftrightarrow f(m)$ entre $1, \ldots, m$, y un subconjunto propio S de los enteros $1, \ldots, n$. Definamos una nueva correspondencia $i \leftrightarrow g(i)$ $[i=1, \ldots, m-1]$ como sigue:

(1) $g(i)=f(i)$ salvo si $f(i)=n$; $g(i)=f(m)$ si $f(i)=n$.

Como $f(i)=n$ para un i a lo más, la correspondencia $i \leftrightarrow g(i)$ será biunívoca entre los enteros $1, \ldots, m-1$, y algunos de los enteros $1, \ldots, n-1$.

Por hipótesis, el conjunto S de todos los enteros $f(i)$ es un subconjunto *propio* del conjunto $1, \ldots, n$; esto significa que alguno de los enteros $1, \ldots, n$ no está en S. Tomemos, pues, el primer número natural $k \leqslant n$ que no esté en S, así que $f(i)$ es menor

(*) Se dice a veces que un conjunto vacío tiene el cardinal *cero*. Preferimos no hacer este convenio (ver la Nota del siguiente § 5).

que k para $i=1, \ldots, m$. Si $k<n$, la definición (1) muestra que nunca será $f(i)=n$, de modo que $g(i)\neq f(m)$. En cualquier caso, los enteros $g(1), \ldots, g(m-1)$ no incluyen a todos los enteros 1, ..., $n-1$, y por ende $i\leftrightarrow g(i)$ es biunívoca entre los enteros 1, ..., $m-1$ y un subconjunto propio de los enteros 1, ..., $n-1$. Ahora, según la hipótesis de inducción, podemos suponer $m-1<n-1$; luego, sumando 1 a los dos miembros, $m<n$, c. q. d.

Corolario 1. *Entre el conjunto 1, ...,* m *y un subconjunto propio del conjunto 1, ...,* n *existirá una correspondencia biunívoca si, y sólo si,* $m \leqslant n$.

Demostración. Si $m\leqslant n$, la correspondencia $1\leftrightarrow 1, \ldots, m\leftrightarrow m$ es del tipo que se ha enunciado. Inversamente, si $i\leftrightarrow f(i)$ es una correspondencia biunívoca entre 1, ..., m y algunos de los enteros 1, ..., n, existirá una correspondencia entre 1, ..., m y un subconjunto propio de 1, ..., n, $n+1$. Luego, por el teorema, $m<n+1$, así que $m\leqslant n$, c. q. d.

Corolario 2. *Si existe una correspondencia biunívoca entre el conjunto 1, ...,* m *y el 1, ...,* n, *debe ser* $m=n$.

Pues, por el Cor. 1, $m\leqslant n$ y $n\leqslant m$, luego $m=n$. Esto demuestra que un mismo conjunto no puede tener como cardinales dos números naturales distintos.

Corolario 3. *Si* S *es un subconjunto propio del conjunto 1, ...,* n, *no existe ninguna correspondencia biunívoca entre éste y* S.

Demostración. Si hubiese tal correspondencia, por el Teorema 1 debiera ser $n<n$, lo que es contradictorio.

El resultado anterior implica inmediatamente el que sigue. Sean S y T dos conjuntos cuyos cardinales respectivos son los números naturales m y n. Entonces $m\leqslant n$ si, y sólo si, existe una correspondencia biunívoca entre S y un subconjunto de T; $m=n$ si, y sólo si, existe una correspondencia biunívoca entre S y todo T.

EJERCICIOS

1. Si un conjunto S tiene cardinal n y si t es un elemento particular de S, mostrar que existe una correspondencia biunívoca entre S y 1, ..., n en la cual t corresponde a n.

2. Si un conjunto S tiene cardinal n, demostrar que la supresión de un elemento de S da un conjunto S' de cardinal $n-1$.
3. Demostrar directamente el Corolario 1, por el método utilizado en la demostración del Teorema 1.
4. Hacer lo mismo para el Corolario 3.

Conjuntos numerables

Un conjunto se llama ordinariamente *finito* cuando sus elementos pueden contarse del modo habitual. Formulemos esta idea de modo más preciso:

Definición. *Un conjunto no vacío S se llama finito cuando su número cardinal es un entero positivo. Un conjunto que no sea vacío ni finito se llama infinito.*

Por ejemplo, el conjunto N de todos los números naturales es infinito (no resulta difícil demostrarlo mediante el Teor. 1). Podemos ahora concebir la idea de que también los conjuntos infinitos tienen sendos números cardinales.

Definición. *Un conjunto S se llama numerable, o se dice que tiene el número cardinal d [en símbolos, $o(S) = d$], si puede ponerse en correspondencia biunívoca con el conjunto de todos los enteros positivos* (*).

Esto equivale a decir que es posible alistar a todos los elementos de S en una sucesión infinita $s_1, s_2, s_3, \ldots, s_n, \ldots$, en la que cada elemento de S aparece una y sólo una vez. Si otro conjunto T está en correspondencia biunívoca con el S, también T será numerable.

Teorema 2 (Paradoja de Galileo). *Cualquier conjunto numerable puede ponerse en correspondencia biunívoca con alguno de sus subconjuntos propios.*

Demostración. Todos los elementos del conjunto dado S pueden escribirse $s_1, s_2, s_3, \ldots$, con los distintos números naturales como subíndices. La correspondencia $s_1 \leftrightarrow s_2, s_2 \leftrightarrow s_3, \ldots, s_n \leftrightarrow s_{n+1}, \ldots$ es biunívoca entre el conjunto S y el que resulta suprimiéndole el elemento s_1.

(*) A veces se usa la letra hebrea alef (subcero) en vez de d.

Puede demostrarse que d (infinito numerable) es el menor número cardinal infinito. En forma precisa, esto es :

TEOREMA 3. *Cualquier conjunto infinito contiene a un subconjunto numerable.*

Sea S tal conjunto y supongamos que se empiezan a contar sus elementos así: $s_1, s_2, s_3, \ldots$ Si quedamos obligados a detenernos, quiere decir que tendremos un número cardinal finito n. De no ser éste el caso, continuaremos indefinidamente (*) obteniendo un subconjunto numerable.

COROLARIO (Dedekind-Peirce). *Un conjunto S es infinito si, y sólo si, puede ser puesto en correspondencia biunívoca con una parte de sí mismo.*

Demostración. Si S es un conjunto finito de número cardinal n, habrá correspondencia biunívoca entre S y $1, \ldots, n$, luego el Corolario 3 al Teorema 1 afirma que S no puede ponerse en correspondencia biunívoca con una parte de sí mismo. Por el contrario, sea S una clase infinita ; S contendrá un subconjunto numerable U, de elementos $u_1, u_2, u_3, \ldots$ La correspondencia que asocia cada u_i de U con el u_{i+1} que le sigue, y cada elemento de S que no esté en U consigo mismo, es biunívoca entre S y uno de sus subconjuntos propios.

En realidad, un sorprendente número de conjuntos importantes resultan ser numerables (es decir, que tienen el cardinal d). He aquí unos ejemplos:

TEOREMA 4. *El conjunto J de todos los enteros es numerable; el conjunto R de todos los números racionales es numerable.*

Demostración. La correspondencia $n \leftrightarrow 2n+1$ $(n = 0, 1, 2, \ldots)$ y $(-n) \leftrightarrow 2n$ $(n = 1, 2, 3, \ldots)$ es biunívoca entre el conjunto $0, -1, +1, -2, +2, \ldots$ de todos los enteros y el conjunto $1, 2, 3, 4, 5, \ldots$ de los enteros naturales. Esto demuestra la primera afirmación.

(*) Si este razonamiento se desarrolla formalmente, se verá que implica un principio básico de la teoría de conjuntos, llamado AXIOMA DE ELECCIÓN. Hemos visto que si los n primeros elementos contados no agotan a S, del subconjunto no vacío de S que resta se ha podido elegir otro elemento s_{n+1}, que ha sido el siguiente de la lista. El axioma de elección afirma, en esencia, que tal conjunto de infinitas elecciones $s_1, s_2, \ldots, s_n, \ldots$ es posible.

Probemos ahora que el conjunto R^+ de todos los números racionales positivos es numerable. Para esto, dispondremos los cocientes de dos números naturales cualesquiera en un cuadro infinito, como el de la fig. 1. Recorriendo ordenadamente los contornos de los cuadrados menores se pueden ordenar todos los cocientes en la siguiente sucesión simple infinita. El primer término es $1/1$; el siguiente de $n/1$ es $1/(n+1)$; el siguiente de m/n es $(m+1)/n$ si $m < n$ y $m/(n-1)$ si $m \geqslant n > 1$. Tachemos ahora de esta sucesión aquellas fracciones que sean iguales a otra previamente enumerada. En la sucesión ordinaria que nos queda, aparece cada número racional positivo una y sólo una vez, de modo que resulta establecida una correspondencia biunívoca $m/n \leftrightarrow k$ entre R^+ y J^+. Pero ésta se puede extender inmediatamente a la correspondencia $m/n \leftrightarrow k$, $0 \leftrightarrow 0$, $(-m/n) \leftrightarrow -k$ entre el conjunto R de todos los números racionales y el conjunto J de todos los enteros. Como J es numerable, también lo será R, c. q. d.

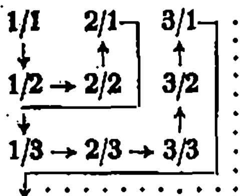


Figura 1

EJERCICIOS

1. Demostrar que el conjunto de todos los enteros múltiplos de 7 es numerable.
2. Demostrar que el conjunto de todos los vectores de un espacio $V_n(R)$ de un número finito de dimensiones sobre el campo de los racionales es numerable.
3. Demostrar directamente que es imposible una correspondencia biunívoca entre el conjunto de todos los números naturales y un conjunto finito.
4. Si $S = T \smile U$, siendo T y U numerables, probar que S es numerable.
5. Si $S = T \smile U$, donde S es numerable y T es finito, demostrar que U es numerable.
6. Demostrar que todo subconjunto de un conjunto numerable es finito o es numerable.
7. Demostrar que los números decimales que terminan en una sucesión de nueves exclusivamente, forman conjunto numerable.
8. Establecer concretamente sendas correspondencias biunívocas entre el conjunto de los enteros y tres subconjuntos propios de él.
9. Demostrar, en la fig. 1, que m/n es el $[(n-1)^2+m]$-ésimo término si $m \leq n$, y el (m^2-n+1)-ésimo si $m > n$.
10. Demostrar que el campo $R(\sqrt{2})$ es numerable (cfr. Cap. II, § 1).
11. Probar que cualquier grupo contiene un subgrupo, numerable o finito.
12. Construir una correspondencia biunívoca entre el campo de los números reales y alguno de sus subconjuntos propios.

13. Demostrar que es numerable el conjunto de todos los números de la forma $r+r'\sqrt{-1}$ (r y r' racionales).

***14.** Demostrar que el anillo $R[x]$ de todos los polinomios con coeficientes racionales es numerable.

3. Otros números cardinales

No todos los conjuntos infinitos son numerables, es decir, no hay sólo un número cardinal infinito. Por ejemplo:

TEOREMA 5 (Cantor). *El conjunto* R^* *de todos los números reales no es numerable.*

Demostración. Vamos a utilizar en ella el llamado «método de la diagonal». Supongamos que existiese una enumeración x_1, x_2, x_3, ... de todos los números reales. Colocando uno sobre otro los desarrollos decimales de todos estos números con las comas en columna, sus cifras decimales formarán un cuadrado como el indicado en la fig. 2. De las cifras que constituyen la diagonal principal deduciremos un número real b, entre 0 y 1, de este modo: llamando a_{nn} la cifra n-ésima de dicha diagonal, como cifra decimal n-ésima del número b pondremos $b_n=a_{nn}-1$ si $a_{nn}\neq 0$, y $b_n=1$ si $a_{nn}=0$. Entonces, $b=0{,}b_1b_2b_3b_4\ldots$ es un número que difiere del x_n en la cifra decimal n-ésima, por lo menos. Así, b no es igual a ningún x_n contra la hipótesis de que la sucesión x_1, x_2, ... comprendía a *todos* los números reales.

$$x_1 = \cdots ,a_{11}a_{12}a_{13}a_{14}\cdots$$
$$x_2 = \cdots ,a_{21}a_{22}a_{23}a_{24}\cdots$$
$$x_3 = \cdots ,a_{31}a_{32}a_{33}a_{34}\cdots$$
$$x_4 = \cdots ,a_{41}a_{42}a_{43}a_{44}\cdots$$

Figura 2

Nota. Esta demostración viene complicada por el hecho de que ciertos números, como $1{,}000=0{,}999\ldots$, pueden tener dos desarrollos decimales distintos, uno terminado en una sucesión de ceros y el otro en una de nueves. La dificultad se salva suponiendo que esta segunda forma de desarrollo (con 9) no se emplea en ningún término de la enumeración original x_1, x_2, ... La construcción de b no da lugar a ninguna cifra 9; por lo tanto, su desarrollo es adecuado para compararlo con el de las x_n.

DEFINICIÓN. *Si un conjunto* S *está en correspondencia biunívoca con el conjunto* R^* *de todos los números reales, se dice que* S *tiene el número cardinal* c *del continuo* [*en fórmula,* o(S)=c].

La mayoría de los conjuntos que intervienen en la Geometría y en el Análisis tienen uno de los números cardinales d o c. Esto puede establecerse caso por caso mediante construcciones adecuadas. Pero a la larga resulta más cómodo demostrar primero un principio general que se debe a F. Bernstein. En la formulación del principio se involucra el concepto general de número cardinal, que ahora vamos a definir.

DEFINICIÓN. *El número cardinal de un conjunto* S *es la colección de todos los conjuntos que pueden ponerse en correspondencia biunívoca con* S (*). *El número cardinal de* S *se indica por* o(S).

Sigue de aquí que dos clases S y T tienen el mismo número cardinal (o son cardinalmente equivalentes) si, y sólo si, existe una correspondencia biunívoca entre ambas. Esto se expresa con la igualdad simbólica $o(S)=o(T)$.

En virtud de la última consideración del §1, el concepto de *desigualdad* entre dos números cardinales puede definirse de modo coherente con la noción ordinaria de desigualdad entre los enteros positivos.

DEFINICIÓN. *Diremos que el conjunto* T *es cardinalmente mayorante del conjunto* S *—y escribiremos* o(S) ⩽ o(T)*— cuando exista una correspondencia biunívoca entre* S *y un subconjunto de* T.

TEOREMA 6 (Bernstein). *Si* o(S) ⩽ o(T) *y* o(T) ⩽ o(S), *debe ser* o(S) = o(T).

Esto quiere decir, en lenguaje ordinario, que si existe una correspondencia biunívoca entre S y una parte de T, y otra entre T y una parte de S, existirá una correspondencia biunívoca entre todo S y todo T. (La recíproca es trivial.)

Demostración. Sea $s \leftrightarrow s\tau$ la correspondencia biunívoca dada entre S y un subconjunto de T, y sea $t \leftrightarrow t\sigma$ la correspondencia biunívoca dada entre T y un subconjunto de S. Cada elemento s de S es la imagen $t\sigma$, a lo más, de un elemento $t=s\sigma^{-1}$ de T; éste (si existe) tendrá a su vez a lo más un *padre* $s'=t\tau^{-1}=s\sigma^{-1}\tau^{-1}$ en S. y así sucesivamente. Retrocediendo de este modo cuanto sea posible

(*) Este concepto se asemeja al de «elemento químico», el cual es igualmente una abstracción que comprende a todos los átomos con una carga nuclear específica.

en los *antepasados* de cada elemento de S (y lo mismo de T), vemos que son posibles tres casos: a) elementos en cuya genealogía se puede retroceder indefinidamente; b) elementos cuya ascendencia termina en un elemento de S; c) elementos cuya ascendencia termina en un elemento de T, miembro el más antiguo de su linaje. En correspondencia a estos casos, dividiremos a S en tres subconjuntos S_a, S_b y S_c, y T en tres subconjuntos T_a, T_b y T_c. Además, la clase que contenga a un elemento de S o de T contendrá asimismo a todos sus *antepasados* y sus *descendientes*.

Ahora bien: σ (¡también τ!) es biunívoca entre S_a y T_a: cada elemento de S_a es la imagen según σ de un elemento de T_a, y sólo de uno, mientras cada elemento t de T_a es el original de un elemento (y sólo de uno) $t\sigma$ de S_a. De modo semejante, τ (¡pero no σ!) es biunívoca entre S_b y T_b, mientras que σ (¡pero no τ!) es biunívoca entre S_c y T_c. Combinando las tres correspondencias biunívocas $S_a \leftrightarrow T_a$, $S_b \leftrightarrow T_b$ y $S_c \leftrightarrow T_c$, obtenemos una correspondencia biunívoca entre todo S y todo T, c. q. d.

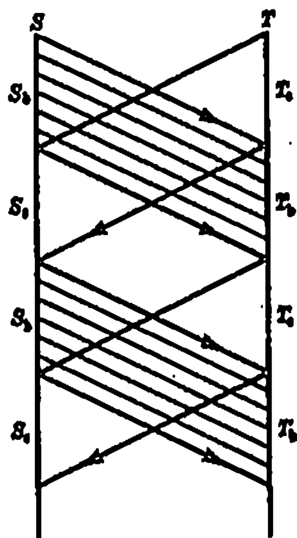


Figura 3

La explicación se ilustra con la figura 3, en la que no aparecen los elementos de genealogía ilimitada. Los conjuntos S y T son los de los puntos de dos rayas verticales, estando τ representado por las flechas que se dirigen hacia la derecha, y σ por las dirigidas hacia la izquierda, mientras el sombreado indica la correspondencia biunívoca entre S_b y T_b.

Teorema 7. *El segmento rectilíneo* $S_1: 0 < x < 1$, *el cuadrado unidad* $S_2: 0 < x, y < 1$ *en el plano, y el cubo unidad* $S_3: 0 < x, y, z < 1$ *en el espacio, tienen el mismo número cardinal.*

Demostración. La correspondencia $x \to e^x = y$ (con inversa $y \to lx$) es biunívoca entre $-\infty < x < +\infty$ y $0 < y < +\infty$; la correspondencia $y \to y/(1+y) = z$ [con inversa $z \to z/(1-z)$] es biuní-

voca entre $0 < y < +\infty$ y $0 < z < 1$. Luego la correspondencia $x \to e^x/(1+e^x) = z$ es biunívoca entre $-\infty < x < +\infty$ y $0 < z < 1$. Esto demuestra la primera afirmación.

Para demostrar la segunda, consideremos la correspondencia

(2) $$(0{,}x_1x_2x_3\ldots,\ 0{,}y_1y_2y_3\ldots) \to 0{,}x_1y_1x_2y_2x_3y_3\ldots$$

entre los pares ordenados de números entre 0 y 1 escritos en forma decimal, y los números entre 0 y 1. Esta correspondencia es biunívoca (aunque no continua) entre el cuadrado S_2 y un subconjunto del segmento S_1 (excluyendo los desarrollos decimales terminados por una sucesión de nueves). Esto demuestra que $o(S_2) \leqslant o(S_1)$. Pero es evidente que $o(S_1) \leqslant o(S_2)$, como lo muestra la correspondencia $x \to (x, 1/2)$; luego, por el Teor. 6, $o(S_2) = o(S_1)$, que es c. Una correspondencia semejante,

(3) $$(0{,}x_1x_2x_3\ldots,\ 0{,}y_1y_2y_3\ldots,\ 0{,}z_1z_2z_3\ldots) \to (0{,}x_1y_1z_1x_2y_2z_2\ldots)$$

demuestra que $o(S_3) \leqslant o(S_1)$, y por consiguiente, como antes, $o(S_3) = c$.

En los ejercicios se ofrecen nuevos ejemplos de conjuntos con número cardinal c.

EJERCICIOS

1. ¿Por qué es trivial el teorema de Bernstein cuando T_0 es vacío? ¿Cómo resulta S_0 en este caso?
2. Expresar los conjuntos S_a, S_b, S_0, T_a, T_b, T_0, en el caso en que S y T son $-1 \leqslant s \leqslant 1/2$ y $-1 \leqslant t \leqslant 1/2$, τ es la correspondencia $s \to s^2$ y σ es la correspondencia $t \to t^2$.
3. El mismo ejercicio si S es el conjunto de los enteros positivos, T el de los enteros no negativos, σ es $s \to s$, τ es $t \to t+1$.
4. Demostrar: cualquier subconjunto de un espacio n-dimensional que contenga un arco continuo tiene cardinal c.
5. Probar que si existe una correspondencia pluriunívoca entre S y la totalidad de otro conjunto T, será $o(T) \leqslant o(S)$.
6. Demostrar que, recíprocamente, si $o(T) \leqslant o(S)$, existirá una correspondencia pluriunívoca entre S y la totalidad de T.
7. ¿Cuáles de las propiedades «reflexiva», «simétrica» y «transitiva» se aplican a la relación de equivalencia cardinal $o(S) = o(T)$? ¿Cuáles se aplican a la relación $o(S) \leqslant o(T)$?
8. Si $o(S) \leqslant o(T)$ y $o(S') = o(S)$, demostrar que $o(S') \leqslant o(T)$.
9. Demostrar: el número de matrices $n \times n$ con elementos cuaternos es c.

10. Establecer explícitamente una correspondencia biunívoca entre el conjunto de todos los números reales entre 0 y 1 y el conjunto de todos los decimales indefinidos $0,a_1a_2a_3...$ (*Sugerencia:* Utilizar el Ejercicio 7 de § 2 y considerar dos decimales indefinidos como iguales cuando sean idénticos, así que $0,100... \neq 0,099...$)
11. El mismo ejercicio para los intervalos $0<x<1$ y $0<x<10$.
12. Sin utilizar el teorema de Bernstein, demostrar directamente:
    a) Que el conjunto de los números reales no negativos tiene cardinal c;
    b) Que el conjunto de todos los números reales positivos que no sean enteros tiene cardinal c.
13. Demostrar: si $S=T\smile U$, con $o(T)=c$ y U numerable, será $o(S)=c$.
14. Demostrar: si $S=T\smile U$, con $o(S)=c$ y T numerable, entonces $o(U)=c$. (*Sugerencia:* Hacer corresponder S con U, previa elección de un subconjunto numerable en U.)

4. Adición y multiplicación de cardinales

Los números cardinales infinitos pueden sumarse y multiplicarse de modo semejante a los ordinarios, conservándose todas las leyes excepto la de simplificación.

Si m y n son enteros positivos, se puede construir un conjunto de número cardinal $m+n$; basta para ello reunir un conjunto S' de número cardinal m (tal como el conjunto 1, 2, ..., m) con otro conjunto S'' disjunto (*) con el anterior y de cardinal n (tal como el $m+1$, $m+2$, ..., $m+n$). La unión $S'\smile S''$ tendrá el cardinal $m+n$. Análogamente, el conjunto de todos los pares (i, j), donde i pertenece al conjunto S' y j al conjunto S'' (que pueden disponerse como los elementos de una matriz $m\times n$), tiene el número cardinal mn. No demostraremos ahora estos hechos tan conocidos; sin embargo, ellos han sido nuestro punto de partida, por cuanto sugieren la ampliación a los cardinales infinitos de la adición y multiplicación ordinarias.

DEFINICIÓN. *Sean α y β dos números cardinales arbitrarios. Llamamos $\alpha+\beta$ al cardinal de aquellos conjuntos que son suma de dos subconjuntos, sin elementos comunes, que tienen α y β como cardinales respectivos; llamamos $\alpha\beta$ al cardinal del conjunto de todos los pares* (x, y), *donde* x *pertenece al conjunto con α elementos, e* y *al conjunto con β elementos.*

(*) Dos conjuntos S' y S'' son *disjuntos*, o sin elementos comunes, cuando su intersección $S'\frown S''$ es nula.

La adición es uniforme. Pues si S y T son suma de dos subconjuntos disjuntos, como S' con S'' y T' con T'' respectivamente, y existe correspondencia biunívoca entre S' y T' de una parte y entre S'' y T'' de otra, se podrán combinar ambas y establecer una correspondencia biunívoca entre todo S y todo T. De un modo semejante resulta uniforme la multiplicación. Por lo demás, otras varias leyes de la aritmética ordinaria se aplican tanto a los números cardinales finitos como a los infinitos (*).

TEOREMA 8. *La adición y multiplicación son conmutativas y asociativas; la multiplicación es distributiva para la adición.*

Demostración. Las leyes conmutativa y asociativa de la adición son consecuencia de las leyes del álgebra booleana. La ley conmutativa de la multiplicación se deduce de que la correspondencia $(x, y) \leftrightarrow (y, x)$ es biunívoca entre el conjunto de todos los pares (x, y) $[x \in S, y \in T]$ y el de todos los pares (y, x) $[y \in T, x \in S]$, cualesquiera que sean S y T. La ley asociativa de la multiplicación se sigue de una evidente correspondencia biunívoca entre el conjunto de todas las ternas $[(x, y), z]$ $[x \in S, y \in T, z \in U]$ y el de las ternas $[x, (y, z)]$ $[x \in S, y \in T, z \in U]$, donde S, T y U son conjuntos arbitrarios. Finalmente, si T y U son disjuntos, $o(S)[o(T)+$ $+o(U)]$ es el número cardinal del conjunto de todos los pares (x, w) $[x \in S, w$ en T o $U]$, mientras que $o(S)o(T)+o(S)o(U)$ es el cardinal de todos los pares (x, y) $[x \in S, y \in T]$ más todos los pares (x, z) $[x \in S, z \in U]$. Hay, pues, una evidente correspondencia biunívoca entre ambos, que demuestra la ley distributiva. Resulta trivial probar que $1 \cdot a = a$, para cualquier cardinal a.

Teorema 9. *Las leyes de simplificación para la adición y la multiplicación no son válidas para números cardinales infinitos.*

Demostración. La demostración del Teorema 2 hace ver que $d = d+1$. Pero esto implica $d+1=(d+1)+1=d+2$, y por otra parte, $1 \neq 2$, de modo que falla la simplificación en la adición. Además, el conjunto J^+ de los enteros positivos es divisible en las dos partes disjuntas de enteros pares o impares, ambas numerables,

(*) Este hecho pierde mucho interés considerando el teorema (que no demostramos) de que la suma o producto de dos números cardinales es simplemente el mayor de los dos. La potenciación transfinita (§ 6) es mucho más interesante.

luego $d+d=d$. Luego, por el Teorema 8, $(1+1)d=1\cdot d$, o sea, $2\cdot d=1\cdot d$, mientras que $2\neq 1$.

Las igualdades $a=a+1$ y $a=a+a$ son válidas para todos los números cardinales infinitos, pero no lo demostraremos.

Resulta de esto que el sistema de los cardinales finitos e infinitos no puede incluirse en ningún sistema en el cual sean posibles la sustracción y la división.

EJERCICIOS

1. Demostrar con detalles (mediante el álgebra booleana) que la adición de números cardinales es conmutativa y asociativa.
2. Demostrar que $a=a+1$ para cualquier cardinal infinito a. (*Sugerencia:* Acudir al Teorema 3.)
3. Demostrar que $d+d+d=ddd=d$. (*Sugerencia:* Ver fig. 1.)
4. a) Si n es un cardinal finito, demostrar que $d+n=d$.
   b) Ver que, semejantemente, es $dn=d$.
5. Demostrar que $c+d=c$ sin utilizar el teorema de Bernstein.
6. Demostrar que $c+c=c\cdot c=c$ sin utilizar § 5.
7. Demostrar $dc=c$.
8. Demostrar la última afirmación del § 4.
9. Dar una demostración del Teorema 8, suponiendo: a) la certeza de lo afirmado en la precedente nota a pie de página; b) que los números cardinales forman un conjunto simplemente ordenado (ver Capítulo XI, § 9, Ejercicio 4).

*10. Para un grupo *numerable* G, consideremos la demostración del Teorema de Lagrange (Cap. VI) sobre el orden de los posibles subgrupos finitos S de G.
   a) Demostrar que la demostración no impone restricciones sobre el orden de S.
   b) Mostrar con ejemplos que pueden existir subgrupos de cualquier orden finito en un G numerable.

* 5. Potenciación

Si S y T son dos conjuntos finitos, de cardinales respectivos $m=o(S)$ y $n=o(T)$, la potencia ordinaria $n^m=o(T)^{o(S)}$ se puede definir como el número de las correspondencias pluriunívocas del conjunto S con cada subconjunto de T. En efecto, una cualquiera de estas correspondencias $x\to y$ está determinada por una función $y=f(x)$ que define para cada argumento x en S un valor y en T. Para contar el número de tales funciones abstractas f, observaremos que el primer elemento x de S tiene precisamente $o(T)$ imágenes y posibles; para cada una de ellas hay otras $o(T)$ elecciones

posibles de la imagen del segundo elemento de S, y así sucesivamente. Así que el número de modos de elegir las imágenes es el producto de $o(T)$ por sí mismo $o(S)$ veces, esto es $o(T)^{o(S)}$.

Esta definición combinatoria de $o(T)^{o(S)}$ puede aplicarse a los números cardinales infinitos.

DEFINICIÓN. *Sean α y β dos números cardinales arbitrarios no nulos. Representaremos por β^α el número de funciones abstractas que transportan un conjunto de α elementos sobre un conjunto de β elementos.*

Omitimos la demostración realmente trivial de la unicidad; es decir, que si $\alpha=\alpha'$ y $\beta=\beta'$, resulta $\beta^\alpha=\beta'^{\alpha'}$.

TEOREMA 10. $c=2^d$.

Demostración. Cada número real entre 0 y 1 tiene un desarrollo diádico distinto $0,x_1x_2x_3\ldots$ (Cap. I, § 12), como sucesión infinita de ceros y unos. Pero el número de estas sucesiones $(x_1, x_2, x_3, \ldots)$ es, por definición, el número de funciones que llevan un dominio numerable (precisamente, el conjunto de los d lugares de orden de esta sucesión) a un dominio de dos elementos (precisamente, el 0 y el 1). Por lo tanto, no hay más de 2^d números reales entre 0 y 1, luego (Teorema 7) $c \leqslant 2^d$.

Por otra parte, cada decimal indefinido compuesto exclusivamente de dos cifras (3 y 5, por ejemplo) representa un número real distinto, luego $2^d \leqslant c$. Ahora, por el Teorema de Bernstein, nos resulta $c=2^d$, c. q. d.

TEOREMA 11. *Para números cardinales cualesquiera α, β y γ, son válidas las siguientes leyes de la potenciación:*

1) $\alpha^\beta\alpha^\gamma=\alpha^{\beta+\gamma}$; 2) $(\alpha\beta)^\gamma=\alpha^\gamma\beta^\gamma$;

3) $(\alpha^\beta)^\gamma=\alpha^{\beta\gamma}$; 4) $\alpha^1=\alpha$ y $1^\alpha=1$.

Demostración. La de las dos partes de 4) es trivial. Para demostrar 1)-3), supongamos que S, T y U son conjuntos de α, β y γ elementos respectivamente, y disjuntos entre sí.

Demostración de 1). Sea V un conjunto del que T y U son subconjuntos complementarios, y consideremos las funciones $h(v)$ que llevan V a S. Por definición, el número de estas funciones es

$\alpha^{\beta+\gamma}$. Por otra parte, cada una de estas funciones está determinada por un par $[f(t), g(u)]$ de funciones independientes, la una llevando de T a S, y la otra de U a S. El número de estos pares es, por definición, $\alpha^\beta \alpha^\gamma$.

Demostración de 2). Consideremos las funciones $h(u)$ que asignan a cada $u \epsilon U$ un par $(s, t) = [f(u), g(u)]$ de valores arbitrarios en S y T, respectivamente. El número de estas funciones es $(\alpha\beta)^\gamma$, por definición. Pero es también el número $\alpha^\gamma\beta^\gamma$ de pares de funciones $f(u)$, $g(u)$, una llevando de U a S, y la otra de U a T.

Demostración de 3). Consideremos las funciones $f(t, u)$ de dos variables $t \epsilon T$, $u \epsilon U$, con valores en S; su número es, por definición, $\alpha^{\beta\gamma}$. Pero también $f(t, u)$ asocia a cada valor fijo de u una función $f_u(t)$ que a cada $t \epsilon T$ le asigna un valor $f_u(t) = f(t, u)$ en S. Recíprocamente, cada correspondencia $u \rightarrow f_u$ define una función $f(t, u) = f_u(t)$ de las variables t y u. Como el número de las f_u es, por definición, α^β, el número de las $f(t, u) = (\alpha^\beta)^\gamma$.

Los teoremas 10-11 nos permiten deducir igualdades en que interviene c, a partir de otras igualdades en que interviene d. Así,

$$c^2 = (2^d)^2 = 2^{2d} = 2^d = c$$
$$2c = 2^1 2^d = 2^{d+1} = 2^d = c$$
$$c^d = (2^d)^d = 2^{d^2} = 2^d = c \qquad \text{(cfr. Teor. 4).}$$

Con estos resultados, el siguiente Ejercicio 1 y el teorema de Bernstein, se obtienen fácilmente fórmulas tales como $d^d = c$, $n^d = c$ para $n > 1$, etc.

Nota. La dificultad de considerar conjuntos vacíos atribuyéndoles el 0, estriba en la ambigüedad de definir el número 0^0 de funciones que transportan un conjunto vacío sobre otro también vacío. ¿Será $0^0 = 0$ o 1? Si $\alpha \neq 0$, $0^\alpha = 0$ y $\alpha^0 = 1$.

Concluiremos con la demostración de una generalización del Teorema 5 de Cantor.

Teorema 12. *Para cualquier número cardinal α, es $\alpha < 2^\alpha$.*

Aclaración. Esta notación significa que $\alpha \leqslant 2^\alpha$, siendo $\alpha \neq 2^\alpha$.

Demostración. Sea S cualquier conjunto de número cardinal α. Por definición, 2^α será el número de funciones $f(x)$, $g(x)$, ... con dominio S y valores 0 y 1. Definiendo, $f_x(y) = 0$ si $x \neq y$ y $f_x(x) = 1$,

tendríamos una correspondencia biunívoca $x \leftrightarrow f_x$ entre S y una clase especial de funciones que transportan a S sobre el conjunto (0, 1). Esto demuestra que $\alpha \leqslant 2^\alpha$.

Recíprocamente, sea una correspondencia biunívoca dada $x \leftrightarrow g_x$ entre S y las funciones de dominio S y valores 0 y 1. Construyendo la nueva función $h(x)$, con $h(x)=0$ si $g_x(x)=1$ y $h(x)=1$ si $g_x(x)=0$, esta $h(x)$ tiene también dominio S y resultante 0 y 1. Sin embargo, $h(x) \neq g_x(x)$ para todo g_x, luego h es distinto de cualquier g_x, luego no existirá ninguna correspondencia biunívoca entre S y el conjunto de *todas* las funciones con dominio S y valores 0 y 1. En símbolos, $\alpha \neq 2^\alpha$.

EJERCICIOS

1. Demostrar que si $\alpha \leqslant \beta$, para todo γ será:

   a) $\alpha+\gamma \leqslant \beta+\gamma$; b) $\alpha\gamma \leqslant \beta\gamma$. c) $\alpha^\gamma \leqslant \beta^\gamma$; d) $\gamma^\alpha \leqslant \gamma^\beta$.

2. Demostrar que $c^c=2^c$. (*Sugerencia:* Ver Ejerc. 7 del § 4.)
3. Si un conjunto S tiene cardinal α, demostrar que el conjunto de todos los posibles subconjuntos de S tiene cardinal 2^α. (*Sugerencia:* Cada subconjunto $T \leqslant S$ determina una llamada *función característica* $f_T(x)$, con $f_T(x)=1$ si $x \in T$, y $f_T(x)=0$ en otro caso.)
4. Demostrar que el número de todos los subconjuntos de un cuadrado es igual al número de todas las funciones reales de variable real.
5. ¿Cuál es el número de conjuntos a) finitos, b) numerables, de números reales?
6. ¿Cuántos conjuntos de números reales de cardinal c existen?

CAPÍTULO XIII

Anillos e ideales

1. Anillos

Los dominios de integridad tratados en el Cap. I son sistemas algebraicos con dos operaciones binarias, adición y multiplicación. Ahora bien, existen sistemas similares en los que la multiplicación no satisface la ley de simplificación o la ley conmutativa, que necesariamente deben cumplirse en aquéllos. A estos sistemas más generales se les llama «anillos». El estudio de los homomorfismos entre anillos conduce a la noción de «ideal», la cual se utiliza con provecho al tratar la teoría de divisibilidad de enteros y en la geometría de curvas y superficies algebraicas.

DEFINICIÓN. *Un anillo* A *es un sistema de elementos tal, que es un grupo abeliano para la operación de adición y es cerrado para una operación binaria de multiplicación, la cual es asociativa y distributiva respecto a la adición. Es decir, para elementos cualesquiera* a, b, c, *del anillo* A, *se tiene*

(1) $$a(bc)=(ab)c, \quad a(b+c)=ab+ac, \quad (a+b)c=ac+bc.$$

En esta definición se incluyen los diversos anillos conmutativos ya conocidos (ver Cap. IV, § 3), tales como el J_m de enteros módulo m y los $A[x]$, $A[x, y]$, de polinomios con coeficientes en un anillo conmutativo dado A. Existen también anillos no conmutativos, como el álgebra de las matrices $n\times n$ y otras álgebras lineales (Cap. VIII, § 6). No es preciso que un anillo tenga elemento unidad; sirva de ejemplo el conjunto de todos los enteros pares $0, \pm 2, \pm 4, \ldots$, que forman un anillo según la anterior definición,

para un elemento unidad. Si A y B son dos anillos, el conjunto $A+B$ de todos los pares (a, b), con a en A y b en B, constituyen un anillo, con las dos operaciones definidas por

$$(2) \qquad (a_1, b_1)+(a_2, b_2)=(a_1+a_2, b_1+b_2),$$
$$(a_1, b_1)(a_2, b_2)=(a_1a_2, b_1b_2).$$

El anillo resultante se llama *suma directa* de A y B y se indica $A+B$.

Muchas de las consideraciones sobre los dominios de integridad han sido hechas independientemente de las leyes de simplificación o conmutativa; son, por lo tanto, aplicables a todos los anillos.

Así, la noción de isomorfismo entre dos dominios de integridad se aplica sin modificación a los anillos. Si se define la noción de *subanillo* por analogía con la de *subdominio*, se puede demostrar que un subconjunto S no vacío de un anillo A es un subanillo si, y sólo si, al estar b y c en S implica que $b-c$ y bc estén en S.

EJERCICIOS

1. a) Demostrar que, en cualquier anillo, $(-a)(-b)=ab$ y $-(-a)=a$.
   b) Si un anillo tiene un elemento unidad 1, demostrar que $(-1)a=-a$, para todo a.
2. Demostrar que la suma directa definida por (2) es un anillo.
3. Demostrar que la suma directa de dos dominios de integridad no es un dominio de integridad.
4. Definir la suma directa de n anillos dados, y probar que es un anillo.
5. Demostrar que la suma directa de dos álgebras lineales sobre un campo F puede considerarse un álgebra sobre F, con una definición conveniente de multiplicación escalar.
6. Definir la locución «subanillo» y demostrar que las condiciones del texto caracterizan a un subanillo.

2. Homomorfismos

Dados dos anillos A y A', se llamará *homomorfismo* entre A y A' (o de A a A') a una correspondencia $a \to aH$ tal, que a cada elemento a de A le corresponda unívocamente un elemento aH de A', y cada elemento a' de A' sea el correspondiente de por lo menos un elemento a de A, $a'=aH$, debiendo verificarse además, para todo a y b de A, que

$$(3) \qquad (a+b)H=aH+bH, \qquad (ab)H=(aH)(bH).$$

Brevemente: un homomorfismo es una correspondencia pluriunívoca, en la cual se conservan sumas y productos.

Un homomorfismo H entre los anillos A y A' es ciertamente un homomorfismo entre el grupo aditivo de A y el de A'. Por lo tanto, H tiene las siguientes propiedades, demostradas para los grupos en §12 del Cap. VI:

(4) $\qquad 0H = 0', \quad (-a)H = -(aH), \quad (a-b)H = aH - bH.$

Aquí, $0'$ es el elemento nulo del anillo A', que es el elemento idéntico del grupo aditivo de A'.

La correspondencia ya conocida $a \to a_m$ de cada entero a con su clase de residuos módulo m, es un homomorfismo entre el anillo J de los enteros y el J_m. Si $f(x)$ es cualquier polinomio con coeficientes en un dominio de integridad D, la correspondencia $f(x) \to f(b)$ establecida «sustituyendo» x por un elemento fijo b de D, es un homomorfismo entre el dominio $D[x]$ y el D, pues las reglas para sumar y multiplicar polinomios en una indeterminada x se aplican ciertamente a las correspondientes expresiones polinómicas en b. Si $R[x]$ es el anillo de polinomios con coeficientes racionales, la correspondencia $f(x) \to f(\sqrt{2})$ representa homomórficamente a $R[x]$ sobre el campo de todos los números $a + b\sqrt{2}$ (ver Cap. II, § 1). La idea de obtener un campo numérico como imagen homomorfa de un anillo de polinomios es importante en el estudio de los números algebraicos (Cap. XVI). La suma directa $A + B$ de dos anillos A y B puede transportarse homomórficamente sobre un sumando B, mediante la correspondencia $(a, b) \to b$; esta correspondencia conserva sumas y productos, por la propia definición (2) de la suma directa.

Para describir explícitamente un homomorfismo particular, debemos preguntarnos, naturalmente, cuándo dos elementos a y b del primer anillo A tienen la misma imagen en el segundo. Por la regla (4), esto sucederá sólo cuando su diferencia tenga como imagen $(a-b)H = 0'$. Por lo tanto, se debe investigar cuál es el conjunto de los elementos de A que tienen como imagen el elemento 0, cero de A'. Por ejemplo, el homomorfismo $J \to J_m$ transporta sobre cero a *todos* los múltiplos km del módulo m. El conjunto de todos estos múltiplos es cerrado para la sustracción, y también para la multiplicación por cualquier entero de J. Análogamente, en el homomorfismo $f(x) \to f(b)$ les corresponde el cero a todos los poli-

nomios divisibles por $x-b$, y no a otros. El conjunto S de tales polinomios es cerrado para la sustracción y para la multiplicación por los elementos de $D[x]$ (estén o no en S). Estos dos ejemplos sugieren la siguiente definición y teorema:

DEFINICIÓN. *Un ideal* C *en un anillo* A *es un subconjunto no vacío de* A *con estas propiedades:*

1) Si c_1 *y* c_2 *están en* C, *también debe estar en* C *la diferencia* c_1-c_2.

2) Si c *está en* C *y* a *en* A, ac *y* ca *están en* C.

TEOREMA 1. *En cualquier homomorfismo* H *de un anillo* A, *el conjunto de elementos que tienen cero por imagen es un ideal en* A.

Para demostrar el Teorema 1 en general, llamemos C al conjunto de todos los elementos c de A para los que $cH=0'$, siendo $0'$ el elemento cero en la imagen A'. En tal caso, para cualquier a en A, $(ac)H=(aH)(cH)=(aH)0'=0'$ y $(ca)H=(cH)(aH)=0'$, lo que demuestra 2). Además, $c_1H=c_2H=0'$ implica, por (4),

$$(c_1-c_2)H=c_1H-c_2H=0'-0'=0',$$

lo que demuestra 1). Este resultado indica que los ideales de un anillo son análogos a los subgrupos normales de un grupo (Cap. VI, Teorema 25).

TEOREMA 2. *Una imagen homomorfa de un anillo* A *está determinada, salvo isomorfismos, por el ideal de elementos representados en cero.*

Demostración. Sean H y K, respectivamente, las representaciones homomórficas de A sobre los anillos A' y A''. Demostraremos que si es $aH=0'$ cuando sea $aK=0''$, y viceversa, entonces serán isomorfos A' y A''. Como es natural, estableceremos que un elemento de A' se corresponda con uno de A'' si, y sólo si, ambos tienen el mismo original en A, esto es,

$$a' \leftrightarrow a'' \quad \text{cuando} \quad aH=a', \quad aK=a'',$$

para algún a. Esta correspondencia es biunívoca, o sea, que bajo ella, cada a' en A' corresponde a una, y sólo una, a'' en A''. Para ver esto, se notará, primero, que cada a' en A' tiene por lo menos

un antecedente a en A, y por lo tanto, le corresponde al menos una $a''=aK$ en A''. En segundo lugar, si $a'\leftrightarrow a''$ y $a'\leftrightarrow b''$, entonces

$$aH=a', \quad aK=a'', \quad bH=a', \quad bK=b'',$$

para ciertos elementos a, b de A, de donde $(a-b)H=a'-a'=0$ implica $0''=(a-b)K=a''-b''$, por hipótesis. La correspondencia también conserva sumas y productos, pues, si $a'\leftrightarrow a''$ y $b'\leftrightarrow b''$, se tendrá

$$a'+b'=(a+b)H\leftrightarrow(a+b)K=a''+b'',$$
$$a'b'=(ab)H\leftrightarrow(ab)K=a''b'',$$

donde a es un original común de a' y a'', y b lo es de b' y b''.

Las dos propiedades 1) y 2) de un ideal tienen varias consecuencias inmediatas. Si un ideal C contiene al elemento c, también, por 1), contendrá al $c-c=0$; por lo tanto, $0-c=-c$ también pertenecerá a C. Aplicando otra vez la propiedad 1), resultará que la suma $c_1+c_2=c_1-(-c_2)$ de dos elementos cualesquiera de C también pertenece a C. Así podemos ver que un subconjunto no vacío C de A es un ideal si, y sólo si, pertenecen a C todas las combinaciones lineales $a_1c_1\pm a_2c_2$ y $c_1a_1\pm c_2a_2$, para c_1 y c_2 en C, y los coeficientes a_1 y a_2 en A. En particular, cualquier ideal de A es un subanillo de A, ya que es cerrado para las dos operaciones de A. El anillo A completo, así como su subanillo (0) que consta del cero solamente, son siempre ideales del anillo. A ambos se les llama *ideales impropios* de A. Los restantes ideales son llamados *propios*.

Correspondientemente, un homomorfismo propio para un anillo A será el que represente sobre cero un ideal propio de A; de modo que no se tratará de un isomorfismo [que representa (0) en $0'$] ni de un homomorfismo que represente a todo el anillo en $0'$.

TEOREMA 8. *Un campo no tiene ninguna imagen homomorfa propia.*

Demostración. Bastará demostrar que un campo, considerado como un anillo, no tiene ideales propios. Sea C un ideal de F distinto del (0), en el cual, pues, estará contenido un elemento $c\neq 0$. Por 2), estará contenido en C el elemento $1=c^{-1}c$ y, también por 2), C contendrá cualquier elemento $a=a\cdot 1$ del campo. Por lo tanto, C es impropio, c. q. d.

Si b es un elemento de un anillo A conmutativo y con unidad, el conjunto (b) de todos los productos xb de b por cualquier x de A es un ideal, pues las propiedades 1) y 2) quedan verificadas. Los ideales del tipo (b) se llaman ideales principales; es claro que (b) es el más pequeño ideal de A que contiene a b.

TEOREMA 4. *En el anillo J de los enteros, todo ideal es principal.*

Esto es simplemente la reafirmación de un corolario del algoritmo para la división de enteros (Cap. I, Teorema 6).

COROLARIO. *Las únicas imágenes propias homomorfas de* J *son los anillos* J_m *de enteros módulo* m.

TEOREMA 5. *En el dominio* F[x] *de los polinomios* f(x) *sobre un campo* F, *todo ideal es principal.*

Este teorema es una nueva formulación del Teor. 8 del Cap. IV, deducido del algoritmo de división para polinomios. Estos resultados no deben producir la falsa impresión de que todo ideal sobre cualquier anillo es principal.

En el anillo $R[x, y]$ de polinomios en dos variables con coeficientes racionales, el conjunto C de todos los polinomios cuyo término independiente es cero, constituye un ideal. Pero no es un ideal principal, pues dos polinomios en x e y pueden pertenecer a C sin que sean por ello múltiplos de un mismo polinomio $f(x, y)$. Así como el ideal C no está engendrado por un solo polinomio $f(x, y)$, todos sus elementos pueden ser representados en la forma $xg(x, y)+yh(x, y)$; esto es decir que el ideal está dado por las combinaciones lineales de *dos* elementos generadores, x e y, con coeficientes polinomios.

Consideremos ahora el ideal engendrado por cualquier conjunto finito de elementos en un anillo A conmutativo y con unidad. Si un ideal C contiene los elementos $c_1, c_2, \ldots, c_m$, debe contener todas las combinaciones lineales $\sum_i x_ic_i$ de estos elementos con coeficientes x_i en A. Pero el conjunto

$$(5)\qquad (c_1, c_2, \ldots, c_m)=[\text{todos los elementos } \textstyle\sum_i x_ic_i \text{ para } x_i \text{ en } A]$$

es él mismo un ideal, pues $\sum_i x_i c_i - \sum_i y_i c_i = \sum_i (x_i - y_i) c_i$, y $a(\sum_i x_i c_i) = \sum_i (a x_i) c_i$, que son las propiedades 1) y 2), características de un ideal. Como A tiene un elemento unidad, cada c_i es uno de los elementos del conjunto (5), a saber, $c_i = 0 \cdot c_1 + \ldots + 0 \cdot c_{i-1} + 1 \cdot c_i + 0 \cdot c_{i+1} + \ldots + 0 \cdot c_m$. Por lo tanto, el conjunto $(c_1, \ldots, c_m)$, definido por (5), es un ideal de A que contiene a todas las c_i y está contenido en cualquier otro ideal que contenga a las c_i. Se le llama el ideal de *base* $c_1, \ldots, c_m$. (No es necesario que los elementos de la base sean linealmente independientes, en lo que *no* se parecen a los de la base de un espacio vectorial.)

En los dominios de integridad más familiares, todos los ideales tienen base finita, pero existen dominios en los que no es éste el caso.

EJERCICIOS

1. ¿Cuáles de las siguientes correspondencias son homomorfismos y por qué? Si la correspondencia es un homomorfismo, identificar al ideal representado en el cero:
   a) $a \to a^2 + 3$, a entero en J;
   b) $a \to 5a$, a entero en J;
   c) $f(x) \to f(\omega)$, $f(x)$ polinomios en $R[x]$, ω raíz cúbica de la unidad;
   d) $f(x) \to f(\omega)$, $f(x)$ polinomios con coeficientes reales;
   e) $f(x, y) \to f(t, t)$, representando $F[x, y]$ en $F[t]$ (x, y, t indeterminadas).
2. Demostrar que cualquier imagen homomorfa de un anillo conmutativo es conmutativa.
3. Si un anillo A tiene un elemento unidad e, demostrar que cualquier imagen homomorfa de A tiene también un elemento unidad.
4. a) Investigar todos los ideales en J_6.
   b) Hallar todas las imágenes homomorfas de J_6.
5. Hacer lo mismo para J_{24}.
6. a) Hallar todos los ideales en J_m para cualquier m.
   b) Hallar todas las imágenes homomorfas de J_m.
7. Hallar todos los ideales en la suma directa de dos campos. Generalizar.
8. Hallar todos los ideales en la suma directa $J+J$, siendo J el anillo de los enteros.
9. Si C_1 y C_2 son ideales en los anillos A_1 y A_2 con unidad, demostrar que $C_1 + C_2$ es un ideal de la suma directa $A_1 + A_2$, y que cualquier ideal de la suma directa tiene esta forma.
10. En un dominio de integridad demostrar que $(a) = (b)$ si, y sólo si, a y b son asociados (Cap. IV, § 5).
11. Si A es un anillo conmutativo con unidad en el que todo ideal es principal, demostrar que dos elementos cualesquiera a y b en A tienen un m. c. d., el cual tiene una expresión $d = ra + sb$.

*12. Sea A un anillo con unidad que contiene a un campo F con el mismo elemento unidad (así, A puede ser un anillo de polinomios sobre F). Demostrar que cualquier imagen propiamente homomorfa de A contiene un campo isomorfo con F.

13. En el anillo $R[x, y]$ de polinomios $f(x, y)=a+b_1x+b_2y+c_1x^2+c_2xy+$ $+c_3y_2+\ldots$, ¿cuáles de los siguientes conjuntos de polinomios son ideales? Si un conjunto es un ideal, hallar su base:

a) Todos los $f(x, y)$ con término constante nulo $(a=0)$;

b) Todos los $f(x, y)$ independientes de x $(b_1=c_1=c_2=\ldots=0)$;

c) Todos los polinomios sin término constante ni lineal $(a=b_1=b_2=0)$;

d) Todos los polinomios sin término cuadrático $(c_1=c_2=c_3=0)$;

e) Todos los polinomios cuyos coeficientes satisfacen la condición $b_1=$ $=-b_2$, $a=0$.

*14. Sea $J_p{}^*$ el anillo de todos los números racionales m/n con denominador primo con p. Demostrar que todo ideal propio en $J_p{}^*$ tiene la forma (p^k) siendo k entero y positivo.

3. Anillo cociente

Cualquier homomorfismo de un anillo tiene un ideal correspondiente, que es el de los elementos representados sobre cero. Ahora, viceversa, dado un ideal, construiremos el correspondiente homomorfismo. Un ideal C en el anillo A es un subgrupo del grupo aditivo de A. Cada elemento a en A pertenece a una clase, llamada *clase de restos* según C, representada por a o por $a'=a+C$, y que consiste en todos los elementos $a+c$ para c variable en C. Dos elementos a_1 y a_2 pertenecen a la misma clase si, y sólo si, su diferencia pertenece al ideal C. Como la adición es conmutativa, C es un subgrupo normal del grupo aditivo A, así que las clases C constituyen un grupo cociente abeliano, en el cual la suma de dos clases es otra clase, obtenida por adición de los elementos representantes, esto es,

(6) $$(a_1+C)+(a_2+C)=(a_1+a_2)+C.$$

Esta suma es independiente de la elección de los elementos a_1 y a_2 en las clases dadas, según demostramos en Cap. VI, § 13.

Para construir el producto de dos clases, escojamos cualquier elemento a_1+c_1 en la primera, y cualquier elemento a_2+c_2 en la segunda. El producto

$$(a_1+c_1)(a_2+c_2)=a_1a_2+(a_1c_2+c_1a_2+c_1c_2)=a_1a_2+c'$$

es siempre un elemento de la clase a_1a_2+C, ya que, por la definición de ideal, la suma $a_1c_2+c_1a_2+c_1c_2$ pertenece al ideal C Por

lo tanto, todos los productos de elementos de la primera clase por elementos de la segunda, pertenecen a una misma clase; este producto de clases es

$$(7) \qquad (a_1+C)(a_2+C)=a_1a_2+C.$$

Las leyes asociativa y distributiva son consecuencia inmediata de las correspondientes leyes en A, así que las clases según C en A, constituyen un anillo.

La correspondencia $a \to a' = a + C$, que hace corresponder a cada elemento de A una de estas clases, es un homomorfismo, por las mismas definiciones (6) y (7) de las operaciones entre clases. En el anillo homomorfo, el elemento cero es la clase $0+C$, así que los elementos de C son representados en el cero. El sumario de todos estos resultados es el siguiente:

TEOREMA 6. *Con las operaciones definidas según (6) y (7), las clases de restos según un ideal* C *de un anillo* A, *forman un anillo, llamado el anillo cociente* (*) A/C. *La correspondencia* a $\to$ a+C, *que hace corresponder a cada elemento de* A *la clase que lo contiene, es un homomorfismo entre* A *y el anillo cociente* A/C, *y los elementos representados sobre cero en este homomorfismo son exactamente los elementos del ideal dado* C.

COROLARIO 1. *Si* A *es conmutativo, también lo es* A/C. *Si* A *tiene elemento unidad, también* A/C *lo tiene.*

La relación entre ideales y homomorfismos es ahora completa. En particular, la unicidad que establece el Teorema 2 puede ser enunciada así:

COROLARIO 2. *Si un homomorfismo* H *representa a* A *en* A′ *y, exactamente, a los elementos de* C *en cero, entonces* A′ *es isomorfo con el anillo cociente* A/C.

El anillo J_m de los enteros módulo m puede ahora expresarse como el anillo cociente $J/(m)$.

Cualquier propiedad de un anillo cociente se reflejará en una propiedad correspondiente de su ideal generador C. Para ilustrar

(*) El anillo A/C es llamado también «anillo de clases de restos» atendiendo a la naturaleza de sus elementos.

este principio, llamemos *máximo* (*) a un ideal $C < A$ cuando los únicos ideales de A que contienen a C son C y el mismo anillo A. Llamaremos *primo* a un ideal P de A cuando cualquier producto ab perteneciente a P tenga al menos un factor, a o b, perteneciente a P. Obsérvese que un número primo p engendra un ideal primo (p) en el anillo J de los enteros, pues un producto ab de dos enteros es múltiplo de p si, y sólo si, uno de los factores es múltiplo de p.

TEOREMA 7. *Si* A *es un anillo conmutativo con elemento unidad, el anillo cociente* A/C *será un dominio de integridad si, y sólo si,* C *es un ideal primo; y será un campo si, y sólo si,* C *es un ideal máximo en* A.

Demostración. El anillo conmutativo A/C es un dominio de integridad si, y sólo si, no tiene divisores de cero (Cap. I, Teorema 1). Esta condición se formulará

(8) $a'b'=0$ implica $a'=0$ o $b'=0$,

donde a' y b' son dos clases de restos según C de los elementos a y b, respectivamente, en A. Ahora bien, una clase a' según C es cero si, y sólo si, a se encuentra en el ideal C, así que la condición (8) puede traducirse por

(9) ab en C implica a en C o b en C.

Resulta, pues, que C cumple exactamente la definición de ideal primo, c. q. d.

Supongamos ahora que C es máximo, y sea b un elemento de A que no pertenezca a C. En tal caso, el conjunto de elementos $c+bx$, para cualquier c en C y cualquier x en A, es un ideal, como es fácil probar. Este ideal contiene a C y contiene a un elemento que no está en C; luego deberá coincidir con todos el anillo A, ya que C es máximo. En particular, la unidad 1 estará en este ideal, así que para alguna a es $1=c+ba$. Traduciéndola en clases, esta ecuación dará $1'=b'a'$. Para cualquier clase $b'=b+C\neq C$ tenemos, pues, una clase inversa $a'=a+C$, lo cual equivale a decir que el anillo conmutativo de clases es un campo. Inversamente, si A/C es un campo, se puede probar que C es máximo.

(*) En vez de «máximo», es muy frecuente llamarle «divisor mínimo».

Como cualquier campo es un dominio de integridad, el Teorema 7 implica que cualquier ideal máximo es primo. La recíproca no es cierta, y un ideal primo puede no ser máximo. Consideremos, por ejemplo, el homomorfismo $f(x, y) \to f(0, y)$, del dominio $F[x, y]$ de todos los polinomios formales en x e y con coeficientes en un campo, al dominio más restringido $F[y]$. El ideal representado en el cero es el ideal principal (x) de todos los polinomios múltiplos de x. Como el anillo imagen $F[y]$ es también un dominio, este ideal (x) es primo (lo cual es fácil de verificar directamente). Pero $F[y]$ no es un campo, luego (x) no puede ser máximo. En efecto, (x) está comprendido en el ideal más amplio (x, y) de todos los polinomios formales con el término constante igual a cero.

EJERCICIOS

1. Demostrar las leyes asociativa y distributiva para la multiplicación de clases residuales.
2. Definamos la congruencia módulo un ideal $C < A$, poniendo $a \equiv b$ (módulo C) para indicar que $a - b$ pertenece a C. Demostrar que las congruencias pueden sumarse y multiplicarse, y mostrar que una clase residual de C consiste en todos los elementos mutuamente congruentes.
3. Demostrar en detalle las dos afirmaciones de Corol. 1, Teor. 6.
4. Hallar todos los ideales primos en el anillo J de los enteros.
5. Hallar todos los ideales primos y todos los ideales máximos en el anillo $F[x]$ de los polinomios sobre F.
6. Demostrar, sin utilizar el Teor. 7, que todo ideal máximo de un dominio de integridad es primo.
7. Hallar un ideal primo que no sea máximo en el dominio $J[x]$ de todos los polinomios con coeficientes enteros.
8. Mostrar que, en el dominio $J[\omega]$ de todos los números $a + b\omega$ (a, b enteros, ω raíz cúbica imaginaria de la unidad), (2) es un ideal primo. Describir $J[\omega]/(2)$.
9. En el anillo de polinomios $R[x, y]$, ¿cuáles de los siguientes ideales son primos y cuáles máximos?

   a) (x^2); c) $(y - 3)$; e) $(x^2 - 1)$;
   b) $(x - 2, y - 3)$; d) $(x^2 + 1)$; f) $(x^2 + 1, y - 3)$
10. Demostrar que si un anillo cociente A/C es un campo, C es un ideal máximo.
11. Hallar anillos que sean isomorfos, respectivamente, a cada anillo cociente A/C de los que siguen:

    a) $A = R[x]$, $C = (x - 2)$;
    b) $A = R[x]$, $C = (x^2 + 1)$;
    c) $A = R[x, y]$, $C = (x, y - 1)$;
    d) $A = J[x]$, $C = (3, x)$;
    e) $A = J_p^*$, $C = (p)$, como en Ejerc. 16 de § 2.

*12. (Segundo teorema del isomorfismo.) Sean $C > D$ dos ideales en un anillo A.
   a) Demostrar que el anillo cociente C/D es un ideal en A/D.
   b) Demostrar que A/C es isomorfo con $(A/D)/(C/D)$. (*Sugerencia:* El producto de dos homomorfismos es un homomorfismo.)

*4. Algebra de ideales

La relación de inclusión entre ideales se asocia estrechamente con la de divisibilidad entre números. En el anillo J de los enteros, $n \mid m$ significa que $m = an$, y por lo tanto, cualquier múltiplo de m es múltiplo de n. Los múltiplos de n constituyen el ideal principal (n), así que lo dicho significa que (m) está incluído en (n). Inversamente, $(m) \leqslant (n)$ significa, en particular, que m está en (n), es decir, que $m = an$. Por lo tanto,

$$(10) \qquad (m) \leqslant (n) \quad \text{si, y sólo si,} \quad n \mid m.$$

Pero ¡cuidado! El número «mayor» corresponde al ideal «menor». Por ejemplo, el ideal (6) de todos los múltiplos de 6 está contenido propiamente en el ideal (2) de todos los números pares.

El m. c. d. y el m. c. m tienen fácil interpretación en la teoría de los ideales. El mínimo común múltiplo m de los enteros n y k es un múltiplo de ambos que es, además, divisor de cualquier otro múltiplo común. El conjunto (m) de todos los múltiplos de m es, pues, el conjunto de todos los múltiplos comunes de n y k; constituye, pues, exactamente el conjunto de elementos comunes a los ideales principales (n) y (k). De modo análogo, en un anillo A, la *intersección* $B \frown C$ de dos ideales B y C es también un ideal. Designando por D otro ideal, el ideal $B \frown C$ tiene estas propiedades:

$$B \frown C \leqslant B \qquad\qquad B \frown C \leqslant C$$
$$D \leqslant B \text{ y } D \leqslant C \quad \text{implica} \quad D \leqslant B \frown C.$$

La intersección es, pues, en el sentido de la teoría de redes, la cota inferior máxima.

Dualmente a la intersección, consideremos la *suma* de dos ideales. Si B y C son ideales de A, se puede comprobar que el conjunto

$$(11) \qquad B + C = [\text{todas las sumas } b + c, \text{ para } b \text{ en } B, c \text{ en } C]$$

es un ideal de A. Como cualquier ideal que contenga a B y C contiene también a todas las sumas $b + c$, dicho ideal $B + C$ contiene

B y C y está contenido en cualquier otro ideal que los contenga a ambos. Así, $B+C$ es la c. s. m. o reunión de B y C, en el sentido de la teoría de redes.

TEOREMA 8. *Los ideales de un anillo* A *bajo la relación de inclusión, constituyen una red, en que la reunión viene dada por la suma* $B+C$ *de (11), y la intersección por* $B \frown C$.

Si dos enteros m y n tienen d por máximo común divisor, el ideal suma $(m)+(n)$ coincide, precisamente, con el ideal principal (d). Pues, por (10), $(d) \geqslant (m)$ y $(d) \geqslant (n)$; mas como d tiene una representación $d=rm+sn$, cualquier ideal que contenga a m y a n deberá contener a d, y, por lo tanto, a (d). Por consiguiente, (d) es la suma de (m) y (n): $(d)=(m)+(n)$.

En general, si los ideales B y C en un anillo conmutativo son engendrados por las bases

(12) $$B=(b_1, \ldots, b_m), \qquad C=(c_1, \ldots, c_n),$$

como en (5), entonces cualquier suma $b+c=\sum_i x_i b_i + \sum_j y_j c_j$ está engendrada por las b_i y c_i, de modo que

(13) $$(b_1, \ldots, b_m)+(c_1, \ldots, c_n)=(b_1, \ldots, b_m, c_1, \ldots, c_n).$$

Esta regla, combinada con la natural transformación de bases, puede utilizarse en el cálculo del m. c. d. de los números enteros. Por ejemplo,

$$(336,\ 270)=(336-270,\ 270)=(66,\ 270)=$$
$$=(66,\ 270-4\times 66)=(66,\ 6)=(6),$$

luego el m. c. d. de 336 y 270 es 6.

En cualquier anillo conmutativo con unidad, se puede también definir el *producto* $B \cdot C$ de dos ideales B y C, de este modo:

(14) $$B\cdot C=[\text{todas las sumas } b_1c_1+\ldots+b_mc_m, \text{ para } b_i \text{ en } B,\ c_j \text{ en } C].$$

Este conjunto es, efectivamente, un ideal; está engendrado por todos los productos bc con un factor en B y otro en C, y es el menor ideal que contiene a tales productos. En particular, el producto de dos ideales principales (b) y (c) es el ideal principal (bc).

Más generalmente: si los ideales B y C están determinados por sendas bases, como en (12), cualquier producto bc tiene la forma

$$bc=(\sum_i x_i b_i)(\sum_j y_j c_j)=\sum_{i,j}(x_i y_j)(b_i c_j).$$

de modo que el ideal producto BC tiene la base

(15) $$BC=(b_1c_1,\ b_1c_2,\ \dots,\ b_mc_{n-1},\ b_mc_n).$$

Estos productos son utilizados en el estudio de los enteros algebraicos (Cap. XIV, § 10).

EJERCICIOS

1. Demostrar detalladamente que $B \frown C$ y $B+C$ son también ideales.
2. Demostrar que el producto BC de (14) es un ideal.
3. Dibujar un diagrama de red para todos los ideales en J_{24}.
4. Si $f(x)$ y $g(x)$ son polinomios sobre un campo y $d(x)$ es su m. c. d., demostrar que $[f(x)]+[g(x)]=[d(x)]$.
5. Calcular por el ideal base el m. c. d. (250, 396) y (8624, 12825).
6. Demostrar que todo ideal en el anillo J de los enteros puede representarse unívocamente como un producto de ideales primos.
7. Demostrar las reglas que siguen para transformar la base de un ideal:

$$(c_1,\ c_2,\ \dots,\ c_m)=(c_1+xc_2,\ c_2,\ \dots,\ c_m),$$
$$(xc_1,\ c_1,\ c_2,\ \dots,\ c_m)=(c_1,\ c_2,\ \dots,\ c_m).$$

8. Simplificar las bases de los siguientes ideales en $R[x,\ y]$:

$$(x^2+y,\ 3y,\ 4x^2+x^2),\qquad (x^2+3xy+y^2,\ 2x^3-y^2,\ x^2+6xy,\ x^2+y^2).$$

9. a) En un anillo conmutativo con unidad, mostrar que $BC \leqslant B \frown C$.
   b) Dar un ejemplo mostrando que el caso $BC < B \frown C$ es posible.
   c) Demostrar que $B(C+D)=BC+BD$.

*10. Demostrar que la red de los ideales en cualquier anillo es modular, en el sentido de Ejerc. 5, Cap. XI, § 9.

11. En un anillo conmutativo A con unidad, designemos por $B:C$ el conjunto de todos los elementos x tales, que xc esté en B siempre que c esté en C.
    a) Si B y C son ideales, demostrar que $B:C$ es también un ideal en A. (Este es el llamado «ideal cociente».)
    b) Demostrar que $(B_1 \frown B_2):C=(B_1:C) \frown (B_2:C)$.
    c) Demostrar que $B:C$ es la c. s. m. de todos los ideales X con $CX \leqslant B$.
12. Demostrar que si un anillo R contiene a los ideales B y C con $B \frown C=0$, $B+C=R$, entonces R es isomorfa con la suma directa de B y C.

*5. Aplicaciones a la geometría algebraica

La noción de ideal es fundamental en las modernas consideraciones sobre la geometría algebraica de las líneas y superficies.

lo apropiado del concepto de ideal para este estudio puede ser ilustrado con un ejemplo sencillo.

El círculo C de radio 2, situado en el plano paralelo al (xy) de cota 2 y con centro en el eje z, se describe analíticamente como el conjunto de puntos (x, y, z) del espacio, que satisfacen al sistema de ecuaciones

(16) $$x^2+y^2-4=0, \quad z-2=0.$$

Con esto, la curva C está definida como intersección de un cilindro circular y de un plano. Pero C puede ser definida, con igual exactitud, como intersección de una esfera y de un plano, según el sistema equivalente

(17) $$x^2+y^2+z^2-8=0, \quad z-2=0.$$

También es posible la descripción por el sistema

(18) $$x^2+y^2-4=0, \quad x^2+y^2-2z=0,$$

que traduce analíticamente la intersección de un cilindro circular con el paraboloide de revolución $x^2+y^2=2z$.

Pero, según esto, la forma más imparcial de describir a C es hacerlo mediante todas las ecuaciones polinómicas a las que satisfacen sus puntos. Mas si $f(x, y, z)$ y $g(x, y, z)$ son dos polinomios cuyos valores resultan idénticamente nulos sobre C, también su suma y diferencia se anularán sobre C. Y lo mismo los productos $a(x, y, z)f(x, y, z)$ por cualquier polinomio $a(x, y, z)$. Esto significa, simplemente, que el conjunto de todos los polinomios idénticamente nulos sobre C es un ideal. Este ideal, y no un par cualquiera de sus elementos, es el que proporciona para C la descripción requerida.

En estas consideraciones vemos que cualquier par de polinomios (16)-(18) es simplemente el generador del ideal de todos los polinomios idénticamente nulos sobre C. Cualquier polinomio obtenido de las ecuaciones (16) por combinación lineal con coeficientes polinómicos, como

(19) $$h(x, y, z)=a(x, y, z)\,(x^2+y^2-4)+b(x, y, z)\,(z-2),$$

pertenece al ideal. Inversamente, puede demostrarse que toda ecuación polinómica $h(x, y, z)=0$ que represente una superficie pa-

sando por la curva C, es de la forma (19). Pero el conjunto de todos los polinomios (19) es simplemente el ideal $(x^2+y^2-4,\ z-2)$ engendrado por los dos polinomios (16) en el anillo $R^*[x, y, z]$ de todas las formas polinómicas en x, y, z con coeficientes en el campo R^* de los números reales. Los polinomios (17) engendran el mismo ideal, pues son combinaciones lineales de los (16), mientras que los (16) pueden obtenerse a su vez como combinación lineal de los (17). El ideal polinómico determinado por la curva C tiene, pues, varias bases:

$$(20)\qquad (x^2+y^2-4,\ z-2)=(x^2+y^2+z^2-8,\ z-2)= \\ =(x^2+y^2-2z,\ z-2).$$

Las ecuaciones paramétricas $x=t$, $y=t^2$, $z=t^3$ determinan una cúbica alabeada C_3. Determinemos el ideal M de los polinomios que se anulen en todo punto de C_3. Estos son los polinomios que se reducen a 0 con la sustitución $x=t$, $y=t^2$, $z=t^3$, esto es, los representados en cero por el homomorfismo

$$(21)\qquad f(x, y, z)\to f(t, t^2, t^3)\qquad (t \text{ es una indeterminada}).$$

Es claro que en todos los puntos de C_3 es $y=x^2$ y $z=x^3$, de modo que $y-x^2$ y $z-x^3$ pertenecen al ideal M. Pero, recíprocamente, observemos que la sustitución $y=y'+x^2$, $z=z'+x^3$ transforma cualquier polinomio $f(x, y, z)$ en otro polinomio $f'(x', y', z')$ y que de esta forma el homomorfismo (21) es

$$(22)\qquad f'(x, y', z')\to f'(t, 0, 0).$$

Esta correspondencia transporta sobre el cero a aquellos términos de f' que contienen y' o z', y no a otros, así que los polinomios representados sobre el cero son, simplemente, las combinaciones lineales $g(x, y, z)y'+h(x, y, z)z'$. Por lo tanto, el ideal M es precisamente el ideal $(y', z')=(y-x^2,\ z-x^3)$ con base $y'=y-x^2$, $z'=z-x^3$. Esto expresa la intersección de dos cilindros. En el estudio de C_3, el anillo cociente $R^*[x, y, z]/M$ juega un importante papel. La representación (22) prueba que este anillo cociente es isomorfo con el anillo de polinomios $R^*[t]$.

La suma de dos ideales tiene una sencilla interpretación geométrica. Por ejemplo, en $R^*[x, y, z]$ el ideal principal $(z-2)$ representa el plano $z=2$, ya que todos los polinomios de la forma

$f(x, y, z)(z-2)$ se anulan cuando, en vez de x, y, z, se ponen las coordenadas de un punto del plano $z=2$. Similarmente, el ideal principal (x^2+y^2-4) define un cilindro de radio 2, cuyo eje es el de las z. La suma de estos dos ideales es $(x^2+y^2-4, z-2)$, de acuerdo con la regla (13). Acabamos de ver que esta suma representa el círculo *intersección* del plano y el cilindro. Y puede demostrarse, efectivamente, que la figura correspondiente a la suma de dos ideales es la intersección de las figuras que corresponden a los sumandos.

En n dimensiones, una *variedad algebraica* (curva, superficie, etcétera) es el conjunto V de todos los puntos $(x_1, \ldots, x_n)$ que satisfacen a un sistema conveniente de ecuaciones polinómicas

$$(23) \qquad f_1(x_1, \ldots, x_n)=0, \ \ldots, \ f_m(x_1, \ldots, x_n)=0.$$

Como antes, el conjunto M_V de los polinomios que se anulan en todos los puntos de V, es un ideal, el cual contiene ciertamente a todas las combinaciones lineales de las $f_n(x_1, \ldots, x_n)$. El *teorema de la base* de Hilbert asegura que M_V tendrá al menos una base *finita*, aunque ésta puede ser distinta de la que sugieren las ecuaciones (23).

Recíprocamente, cualquier ideal C del anillo de polinomios $R^*[x_1, \ldots, x_n]$ determina una figura correspondiente, que consiste en todos los puntos $(a_1, \ldots, a_n)$ del espacio n-dimensional tales, que $f(a_1, \ldots, a_n)=0$ para todo polinomio $f \in C$. El teorema de la base, de Hilbert, asegura que C tiene una base finita $f_1, \ldots, f_m$, así que la correspondiente figura V es precisamente una variedad algebraica. Sin embargo, el ideal M_V de esta variedad puede ser más amplio que el ideal dado C (ver el siguiente Ejerc. 3).

EJERCICIOS

1. Hallar el ideal correspondiente a la curva de ecuaciones paramétricas $x=t+1$, $y=t^3$, $z=t^4+t^2$.
2. Demostrar que un ideal $(ax+by+cz+d,\ a'x+b'y+c'z+d')$ engendrado por dos polinomios lineales, linealmente independientes, determina una recta.
3. a) Demostrar que los ideales (x, y) y (x^2, xy, y^2) en $R^*[x, y, z]$ determinan la misma variedad algebraica.

   b) Mostrar que la correspondencia entre los ideales de polinomios y las variedades algebraicas que determinan, no es biunívoca.

   c) Demostrar que cualquier ideal y su cuadrado determinan el mismo lugar.

Mostrar con detalle que el conjunto de polinomios en $R^*[x_1, \ldots, x_n]$ que se anulan idénticamente sobre un lugar C, es un ideal.

a) ¿Cuál es el lugar determinado por $xy=0$ en el espacio tridimensional?

b) Probar que el lugar determinado por el *producto* de dos ideales principales es la *unión* de los lugares determinados por cada ideal separadamente.

c) Generalizarlo a ideales arbitrarios. (*Sugerencia:* Si un punto del lugar del producto no está en el lugar determinado por el primer factor, deberá anular a uno por lo menos de los polinomios del primer ideal.)

d) ¿Cuál es el lugar determinado por la intersección de dos ideales?

a) Calcular la inversa de la transformación «birracional» T: $x'=x$, $y'=y-x^2$, $z'=y-x^3$.

b) Demostrar que el conjunto de todas las substituciones $x'=x$, $y'=y+p(x)$, $z'=z+q(x, y)$ es un grupo.

c) Mostrar que cada una de estas substituciones inducen un *automorfismo* sobre el anillo $R^*[x, y, z]$.

a) Si C es un ideal en un anillo A, el *radical* de C es el conjunto $\sqrt{C}$ de todos los x en A con alguna potencia x^m en C. Demostrar que $\sqrt{C}$ es un ideal.

b) Si C es un ideal en el anillo de polinomios del texto, V el lugar correspondiente, demostrar que M_V contiene a $\sqrt{C}$. (Un teorema de Hilbert asegura que $M_V = \sqrt{C}$.)

Ideales en las álgebras lineales

En un anillo no conmutativo cabe considerar ideales por la ierda y por la derecha. Un *ideal por la izquierda* L en un anies un subconjunto de A tal, que $x-y$ y ax están en L siemque lo estén x e y y a sea de A. Un *ideal por la derecha* se ne de modo análogo. Y cuando un ideal lo sea en el sentido que os considerado hasta ahora, se le llama ideal *por ambos lados* mplemente «ideal». Por ejemplo, en el anillo M_2 de todas las rices 2×2, las matrices en las que la primera columna es totalte de ceros forman un ideal por la izquierda, pero no por la cha.

Estas ideas se pueden aplicar útilmente a un álgebra lineal A un elemento unidad 1. En este caso, cualquier ideal por la ierda L o por la derecha R es también cerrado para la multiación por un escalar, es decir, que si ξ es un elemento de L es un escalar, L contendrá también a $c\xi$, pues, en efecto, $(c \cdot 1)\xi$ es el producto de un elemento en L por un elemento de A. Por lo tanto, si A se considera como un espacio lineal

sobre el campo de coeficientes, cualquier ideal por la izquierda (o por la derecha) de A es un subespacio.

Este resultado puede enunciarse de otro modo como sigue. Una *subálgebra* de un álgebra lineal A es un subespacio S de A cerrado para la multiplicación. En fórmulas,

(24) $\xi \in S$ y $\eta \in S$ implican $c\xi + d\eta \in S$ y $\xi\eta \in S$.

Una subálgebra de A se llama *invariante por la izquierda* si cuando contiene a un elemento ξ, contiene también a todos sus múltiplos por la izquierda $\alpha\xi$ $(\alpha \in A)$; las subálgebras *invariantes por la derecha* y las *invariantes* se definirán análogamente a como se hizo, respectivamente, con los ideales por la derecha y con los ideales (por ambos lados).

Hemos visto, pues, que en el caso de las álgebras lineales con elemento unidad, esta terminología adicional es superflua:

TEOREMA 9. *Las subálgebras invariantes por la izquierda, invariantes por la derecha e invariantes (por ambos lados) de un álgebra lineal con elemento unidad, son respectivamente equivalentes a sus ideales por la izquierda, ideales por la derecha e ideales (por ambos lados).*

Un álgebra lineal se llama *simple* si no tiene ideales (por ambos lados) propios. Por lo tanto, un álgebra simple no tiene imágenes homomorfas propias.

TEOREMA 10. *El álgebra de todas las matrices* n × n *sobre un campo es simple.*

Demostración. Esta álgebra M_n tiene como base las n^2 matrices E_{ij}, cuyo elemento en la posición (i, j) es 1, siendo nulos los restantes. Un ideal propio B en M_n debería contener al menos una matriz no nula $A = \sum_{i} a_{ij}E_{ij}$, con algún coeficiente $a_{rs} \neq 0$. Pero entonces, la matriz

$$(a_{rs})^{-1}E_{kr}AE_{sk} = (a_{rs})^{-1}\sum_{i,j} E_{kr}E_{ij}E_{sk}a_{ij} = E_{kk}$$

estará también en B. Por lo tanto, también pertenecerá a B la matriz identidad $I = \sum_{k} E_{kk}$, luego B coincidirá con M_n y no se tratará de una subálgebra propia, c. q. d.

Wedderburn (1908) demostró un célebre recíproco del Teorema 10. Este recíproco afirma en particular que toda álgebra simple sobre el campo C de los números complejos es isomorfa con el álgebra de todas las matrices $n \times n$ sobre C. Para tratar el caso general, se necesitan los conceptos de álgebra de división (Cap. VIII, §6) y de álgebra de matrices cuyos coeficientes a_{ij} pertenecen a un álgebra de división.

Para sumar o multiplicar dos matrices $n \times n$ con coeficientes en un álgebra de división D, se aplican simplemente las reglas ordinarias :

$$(25) \qquad \begin{gathered} \| a_{ij} \| + \| b_{ij} \| = \| a_{ij} + b_{ij} \|, \qquad c \| a_{ij} \| = \| c a_{ij} \|, \\ \| a_{ij} \| \cdot \| b_{ij} \| = \| \sum_{k=1}^{n} a_{ik} b_{kj} \|. \end{gathered}$$

El resultado de Wedderburn es que si F es cualquier campo, la más general álgebra simple A sobre F se obtiene como sigue. Tomemos cualquier álgebra de división D sobre F y cualquier entero positivo n. Entonces A consta de todas las matrices $n \times n$ con coeficientes en D. En el caso particular de ser F el campo C de los números complejos, el teorema fundamental del álgebra se puede utilizar para demostrar que la única álgebra de división sobre C coincide con C.

EJERCICIOS

1. Demostrar que cualquier álgebra de división es simple.
2. Hallar todos los ideales por la derecha en un álgebra de división.
3. Discutir el álgebra de los ideales por la izquierda en un anillo, en el sentido de §4.
4. Demostrar que cualquier anillo cociente de un álgebra lineal sobre F es también un álgebra lineal.
5. a) Si S es un subespacio del espacio vectorial $V_n(F)$, demostrar que el conjunto de todas las matrices con filas en S es un ideal por la izquierda de $M_n(F)$.
   * b) Demostrar que todo ideal por la izquierda C de $M_n(F)$ es uno de los descritos en la parte a). (*Sugerencia:* Ver que cualquier fila de una matriz de C es la primera fila de una matriz en C que tiene iguales a 0 todas las filas restantes. Utilizar los métodos del Capítulo X.)

7. Característica de un dominio de integridad

Resulta difícil la clasificación de todos los anillos, pero en los dominios de integridad se llega a resultados satisfactorios exami-

nando los subgrupos cíclicos de D, considerando D como un grupo aditivo (*). Para la potencia m-ésima de un elemento de este grupo emplearemos sistemáticamente la notación $m \times a$. Así, si m es un entero positivo,

$$m \times a = a + \ldots + a \quad (m \text{ sumandos}), \tag{26}$$

si $m=0$, $0 \times a = 0$, mientras que si $m = -n$ es negativo,

$$-n \times a = n \times (-a) = (-a) + \ldots + (-a) \quad (n \text{ sumandos}). \tag{27}$$

Para destacar que $m \times a$ *no* es el producto de a por otro elemento de D, llamaremos a $m \times a$ el m-ésimo *múltiplo natural* de a; m es simplemente un entero ordinario, mientras que los elementos de D pueden no ser números.

Los múltiplos naturales de los elementos del dominio D cumplen todas las propiedades que han sido demostradas, con la notación multiplicativa, para las potencias de un grupo conmutativo; por lo tanto,

$$(m \times a) + (n \times a) = (m+n) \times a, \quad m \times (n \times a) = (mn) \times a, \tag{28}$$

y

$$m \times (a+b) = m \times a + m \times b, \quad m \times (-a) = (-m) \times a. \tag{29}$$

Veamos ahora otras propiedades, consecuencia de la ley distributiva. La ley general distributiva (ver Cap. I, § 5) es

$$(a+a+\ldots+a)b = ab + ab + \ldots + ab \quad (m \text{ sumandos}).$$

Esta fórmula, con la notación de los múltiplos naturales, se escribe así:

$$(m \times a)b = m \times (ab) = a(m \times b). \tag{30}$$

Lo mismo es válido para $m=0$ y para m negativo, pues si $m = -n$, la definición (27) da

$$(-n) \times ab = n \times (-ab) = [n \times (-a)]b = [(-n) \times a]b.$$

(*) Representar D en notación aditiva supone considerar a 0 como elemento idéntico, $(-a)$ como inverso de a, y $a + \ldots + a$ como potencia m-ésima de a.

La regla $(a+\ldots+a)(b+\ldots+b)=ab+\ldots+ab$ es otra ley distributiva general. Se puede formular de nuevo así:

$$(m\times a)\,(n\times b)=(mn)\times(ab),$$

que también vale para los enteros m y n positivos, negativos o nulo. Apliquemos la idea de múltiplos naturales al dominio J_p de los enteros módulo p, con p primo. Como cualquier entero a repetido p veces como sumando da una suma divisible por p, cualquier elemento a en J_p tendrá la propiedad característica de ser

$$p\times a=a+a+\ldots+a=0 \quad (p \text{ sumandos}).$$

Esto significa que la «potencia» p-ésima de a en el grupo aditivo J_p es siempre el elemento 0 del grupo. El orden de a en el grupo es, por lo tanto, p o un divisor de p. Como p es primo, no puede tener divisores propios, y resulta que cualquier elemento no nulo de J_p tiene el orden aditivo p. Por otra parte, en el dominio de integridad J de todos los enteros, la adición repetida de un entero con él mismo no puede ser cero, luego en dicho dominio todo entero no nulo tiene orden aditivo infinito. Estos dos hechos pueden generalizarse.

TEOREMA 11. *En el grupo aditivo de un dominio de integridad, todos los elementos no nulos tienen el mismo orden.*

Por definición, el orden de a es el menor entero m tal, que el múltiplo natural $m\times a$ es cero. Luego nos bastará probar que, dados $a\neq 0$ y $b\neq 0$, es $m\times a=0$ si, y sólo si, $m\times b=0$. Supongamos que $m\times a=0$. Entonces $(m\times a)b=0$, y por (80),

$$(m\times b)a=m\times(ab)=(m\times a)b=0.$$

El factor $a\neq 0$ puede suprimirse por la ley de simplificación y queda $m\times b=0$. Esto demuestra el teorema.

DEFINICIÓN. *Se llama característica de un dominio de integridad al orden común de sus elementos no nulos en el grupo aditivo del dominio.*

Como el elemento unidad e (*) no es nulo, se puede decir que la característica del dominio es el menor entero m tal, que $m\times e=0$.

(*) En este razonamiento representamos por e la unidad del dominio, distinguiéndole así del entero positivo 1.

Si $m \times e$ es siempre distinto de cero, para cualquier positivo m, se dirá que la característica es ∞ (*). El dominio J de todos los enteros tiene característica ∞, mientras que el dominio J_p tiene característica p. Estos son los únicos casos posibles.

TEOREMA 12. *La característica de un dominio de integridad, o bien es ∞, o bien es un número primo* p.

Para demostrarlo, supóngase que, al contrario, algún dominio D tuviera característica compuesta, como $m = rs$. Entonces, por (31),

$$0 = m \times e = (rs) \times e = (r \times e) \cdot (s \times e).$$

Por la ley de simplificación, o bien será $r \times e = 0$ o bien $s \times e = 0$. Luego la característica deberá ser un divisor de r o de s, pero no m, como habíamos supuesto.

Este resultado nos dice que un dominio D o bien tiene característica prima p (**), en cuyo caso vale (32), o bien característica ∞, en cuyo caso $a + a + \ldots + a = 0$ implica $a = 0$. En el estudio de las formas cuadráticas (Cap. IX, Teor. 13), nos hemos encontrado con algunos teoremas válidos en los campos de característica distinta de 2.

El dominio fundamental de característica p es el J_p, dándole a esta afirmación el siguiente sentido:

TEOREMA 13. *En cualquier dominio* D *de característica* p, *el subgrupo aditivo* S *engendrado por el elemento unidad es un subdominio isomorfo con el dominio* J_p *de los enteros módulo* p.

Demostración. El subgrupo cíclico en cuestión, S, consiste en todos los múltiplos naturales distintos $1 \times e$, $2 \times e$, ..., $(p-1) \times e$, $p \times e = 0$ del elemento e unidad de D. Por (31) es $(m \times e)(n \times e) = (mn) \times e$, así que S es multiplicativamente cerrado (por otra parte, contiene elemento unidad e y es cerrado para la adición y sustracción), luego S es un subdominio de D. Este subdominio consta exactamente de p elementos distintos y puede ser puesto en correspondencia con el dominio J_p de los enteros módulo p por la regla $m \times e \leftrightarrow m_p$, donde m_p es la clase de m en el dominio J_p de las

(*) Muchos autores dicen «característica 0» o «sin característica», en vez de «característica ∞».

(**) Los campos de característica $p \neq \infty$ se llaman a veces «campos modulares».

clases de restos, módulo p. Bajo esta regla, cada elemento de J_p es el correspondiente de un elemento de S, e inversamente.

Para probar que la correspondencia es biunívoca, debemos demostrar que $m \times e = n \times e$ si, y sólo si, $m_p = n_p$. Pero $m_p = n_p$ significa que $m \equiv n$ (mód. p), o sea que $p \mid (m - n)$, y esto, a su vez, equivale sucesivamente a $(m - n) \times e = 0$, $m \times e - n \times e = 0$, $m \times e = n \times e$.

Para probar que esta correspondencia es un isomorfismo $J_p \leftrightarrow S$, será suficiente comparar las fórmulas de adición y multiplicación. Pero las reglas para sumar y multiplicar las clases residuales en J_p son exactamente paralelas a las fórmulas (28) y (31),

$$(m \times e) + (n \times e) = (m + n) \times e, \quad m_p + n_p = (m + n)_p,$$

$$(m \times e)(n \times e) = (mn) \times e, \quad m_p n_p = (mn)_p.$$

En el caso de característica ∞, la correspondencia $m \times e \leftrightarrow m$ da un isomorfismo similar; así se puede demostrar el siguiente resultado:

TEOREMA 14. *En cualquier dominio D de característica ∞, el subgrupo aditivo engendrado por el elemento unidad es un subdominio isomorfo con el dominio J de los enteros.*

Este teorema (cfr. Teor. 20, Cap. I) es una seria limitación de la aparente gran generalidad de la noción de dominio. De acuerdo con la definición, un dominio puede ser cualquier conjunto de elementos cualesquiera entre los cuales están definidas la adición y la multiplicación con propiedades determinadas. Hemos encontrado varios ejemplos de dominios de característica ∞ (el dominio de todos los números racionales, el de todos los números reales, el de todos los números complejos, el de todas las formas polinómicas con coeficientes racionales, etc.). Cada uno de estos dominios contiene al de los números enteros como un subdominio. Ahora, el Teorema 14 nos enseña que esta situación es inevitable, pues cada dominio de característica ∞ debe contener también a todos los enteros o a alguna cosa exactamente análoga a ellos. Si no se insiste en la diferenciación entre dominios diversos pero isomorfos, se puede decir que cualquier dominio puede ser obtenido añadiendo los elementos convenientes a J o a algún J_p.

En un dominio de característica ∞, los múltiplos $m \times a$ pueden ser considerados como productos ordinarios, pues por (30) $m \times a = (m \times e)a$, donde $m \times e$ es el elemento de D que corresponde al entero m en el isomorfismo utilizado en el Teorema 14. En los dominios de característica p, los múltiplos deben ser tratados con más precaución, porque, en tal caso, $m \times a = 0$ no implica que $m = 0$ o $a = 0$. Por ejemplo, el sencillo desarrollo

$$(a+b)^2 = a^2 + ab + ba + b^2 = a^2 + 2 \times (ab) + b^2$$

tiene un término medio que es, propiamente hablando, un múltiplo $2 \times (ab)$ (y no un producto $2ab$). Por el mismo motivo, en la fórmula ordinaria del desarrollo de la potencia del binomio los coeficientes binómicos intervienen como *múltiplos naturales*. En cualquier dominio de integridad se puede establecer la fórmula del binomio para un exponente natural n, en la forma

$$(33) \qquad (a+b)^n = a^n + C_1^n \times (a^{n-1}b) + C_2^n \times (a^{n-2}b^2) + \dots + C_n^n \times b^n,$$

donde los coeficientes $C_i^n = \binom{n}{i}$ son enteros dados por las fórmulas

$$(34) \qquad C_i^n = [n!]/[(n-i)!\,i!], \qquad i = 0, 1, \dots, n,$$

donde $n! = n(n-1) \dots 3 \cdot 2 \cdot 1$.

Apliquemos esta fórmula para $n = p$ y dos elementos a, b, en un dominio de integridad D de característica p. Entonces $(a+b)^p$ es

$$a^p + p \times (a^{p-1}b) + \frac{p(p-1)}{2} \times (a^{p-2}b^2) + \\ + \frac{p(p-1)(p-2)}{1 \cdot 2 \cdot 3} \times (a^{p-3}b^3) + \dots + b^p.$$

En los coeficientes intermedios, el numerador admite el factor primo p, mientras que los factores del denominador, por ser menores que p, son primos con él. Por lo tanto, cada coeficiente binómico es un entero múltiplo de p. Pero como el dominio tiene característica p, el producto por p de los múltiplos de sus elementos debe ser cero; el desarrollo resulta, por tanto,

$$(35) \qquad (a+b)^p = a^p + b^p.$$

También se demuestra que en un dominio de característica p es

$$(a-b)^p=a^p-b^p. \tag{36}$$

Si p es impar, esto sigue inmediatamente de (35) porque $(a-b)^p=[a+(-e)b]^p=a^p+(-e)^pb^p$ y $(-e)^p=-e$. Si p es un primo par, es $p=2$ y $(-e)^p=+e$, pero también $+e=-e$, puesto que $e+e=0$.

El resultado (35) establece que la correspondencia $a\rightarrow a^p$ transforma sumas en sumas. También conserva los productos, pues la igualdad $(ab)^p=a^pb^p$ resulta válida en todo dominio de integridad. Esta correspondencia representa al dominio dado D en el conjunto D^p de todas las potencias p-ésimas posibles de los elementos de D. La correspondencia entre D y D^p es biunívoca, porque $a^p=b^p$ implica por (36) que $(a-b)^p=0$, y por lo tanto, $a=b$. Además, el conjunto D^p es un subdominio de D, ya que es cerrado para la adición, sustracción y multiplicación. Estas consideraciones demuestran el siguiente resultado:

TEOREMA 15. *La correspondencia* $a\rightarrow a^p$ *representa isomórficamente cualquier dominio* D *de característica* p *sobre el subdominio* D^p *de las respectivas potencias* p-*ésimas de todos los elementos de* D.

Por ejemplo, si D es el dominio J_p de los enteros módulo p, el isomorfismo $a\rightarrow a^p$ es simplemente la correspondencia idéntica, de acuerdo con el teorema de Fermat (Cap. VI, § 9).

EJERCICIOS

1. Demostrar que el múltiplo natural $m\times a$ puede definirse para m positivo por las fórmulas de recurrencia $1\times a=a$; $(m+1)\times a=m\times a+a$.
2. Demostrar por inducción las reglas (28) y (30) para múltiplos naturales.
3. Demostrar con detalle el Teor. 14.
4. ¿Qué resulta del Teor. 11 para el conjunto de los enteros mód. m., si m es compuesto?
5. ¿Qué parte de la teoría de características puede aplicarse a los anillos conmutativos que no sean un dominio de integridad?
6. a) Demostrar, en Teorema 15, que si D es finito, es $D^p=D$.
   b) Demostrar que si $D=J_p[x]$ es $D^p<D$.
   * c) Demostrar que todo campo finito, distinto de los campos «primos» J_p, tiene un automorfismo distinto de la identidad.

7. ¿Puede decirse algo sobre la característica de un dominio de integridad ordenado?

Característica de un campo

Puesto que un campo es un dominio de integridad en el cual es posible la división (excepto por cero), las anteriores consideraciones sobre la característica se aplican también a los campos. Si un campo F tiene característica p, entonces, por el Teorema 13, el subgrupo aditivo de F engendrado por su elemento unidad es un subcampo, y es isomorfo con el campo finito de los enteros módulo p. Si un campo F tiene característica ∞, entonces, por el Teorema 14, el subgrupo engendrado por el elemento unidad e consta de todos los múltiplos $m \times e$, así que el subcampo engendrado por e se compone de todos los cocientes $(m \times e)/(n \times e)$, con $n \neq 0$. Este subcampo es el campo de cocientes del subdominio de todos los múltiplos $m \times e$. Éste, por el Teor. 6 de Cap. II, es isomorfo con el campo de los números racionales, el cual es el campo de cocientes del dominio de enteros $m \leftrightarrow m \times e$. Concretamente, la correspondencia $(m \times e)/(n \times e) \leftrightarrow m/n$ es un isomorfismo entre el campo engendrado por e y el campo de los números racionales. Esto demuestra el siguiente resultado (cfr. Teor. 13, Cap. II):

TEOREMA 16. *En un campo de característica ∞, el subcampo engendrado por el elemento unidad es isomorfo con el campo* R *de todos los números racionales.*

El isomorfismo $(m \times e)/(n \times e) \leftrightarrow m/n$ conserva las cuatro operaciones racionales en un tal campo F. Tratándose de un campo particular F, es, pues, posible (y conveniente) identificar cada elemento de la forma $(m \times e)/(n \times e)$ con el correspondiente número racional m/n. Con este convenio, todo campo de característica ∞ viene a contener todos los números racionales m/n, con $n \neq 0$. Por un convenio similar, cualquier campo de característica p viene a contener el campo J_p. En este sentido, cualquier campo es una extensión de uno de los campos mínimos (también llamados campos *primos*) R y J_p. Por lo tanto, es natural comenzar la clasificación sistemática de los campos, considerando los modos de ampliar un campo dado. De esto nos ocuparemos en el próximo capítulo.

EJERCICIOS

1. Sea F_4 un campo con, precisamente, cuatro elementos.
   a) Demostrar que F_4 tiene característica 2;
   b) Demostrar que los dos elementos de F_4 ajenos al subcampo primo J_2 satisfacen a $x^2=x+1$;
   c) Utilizando este hecho, demostrar que F_4 es isomorfo con el campo $J[\omega]/(2)$ del Ejerc. 8, § 3.
2. Hallar todos los automorfismos del campo F_4 del Ejerc. 1.
3. Demostrar que la fórmula corriente para la solución de una ecuación cuadrática se aplica a cualquier campo de característica distinta de 2.
4. ¿Sobre qué campos es válida la fórmula usual para resolver una ecuación cúbica?

CAPITULO XIV

Campos de números algebraicos

1. Ampliaciones algebraicas y trascendentes de un campo

El estudio de las ecuaciones polinómicas puede desarrollarse ahora desde un punto de vista más general. En vez de considerar aisladamente cada raíz de una ecuación tal como $x^3-2=0$, es preferible considerar todo el conjunto de números que pueden ser obtenidos a partir de esta raíz, con operaciones racionales. Este conjunto de números es un campo, que aparece como una extensión o ampliación del campo original (esto es, del campo de los números racionales). Con esta idea se considerarán sistemáticamente las posibles ampliaciones del campo R de los números racionales. Es igualmente fácil considerar las ampliaciones de *cualquier* campo F. Este aumento de generalidad tiene diversas ventajas. Cuando se le aplica a campos F de característica p, hace posible una clasificación sistemática de todos los campos finitos (Capítulo XV, § 6); cuando se utiliza para un campo F de formas racionales, es la base de la teoría de las funciones algebraicas y sus integrales.

Por *ampliación* K de un campo F entendemos, simplemente, todo campo que contiene a F como subcampo. Tal ampliación puede ser engendrada a partir de F por ciertos elementos de la misma; por ejemplo, los números complejos $a+bi$ son engendrados por los reales y el solo número complejo i, mientras que el campo $R(x)$ de todas las formas racionales (con coeficientes racionales) en una indeterminada x, es engendrado por el campo R y el

elemento x. Un mismo campo puede ser engendrado de muy diversas maneras. Por ejemplo, el campo $R(\sqrt{2})$ es engendrado por una raíz $\sqrt{2}$ de la ecuación $x^2-2=0$, y consisten todos los números reales $a+b\sqrt{2}$ con coeficientes a y b racionales (ver el ejemplo de Cap. II, § 1). La ecuación $x^2+4x+2=0$, que es distinta de la anterior, tiene la raíz $-2+\sqrt{2}$, la cual engendra el mismo campo $R(\sqrt{2})$, pues todo número de este campo puede ser expresado por el número generador como sigue:

$$a+b\sqrt{2}=(a+2b)+b(-2+\sqrt{2}).$$

Aplicando a nuestra ecuación el procedimiento usual de completar el cuadrado, quedará $x^2+4x+2=(x+2)^2-2=0$, así que obtenemos de nuevo la ecuación $y^2-2=0$ con una raíz que engendra el mismo campo. El efectuar una transformación de variables para simplificar una ecuación, puede tomarse, pues, como sinónimo de elegir un nuevo generador para el correspondiente campo.

Vamos ahora a definir con toda generalidad el subcampo *engendrado* por un elemento dado. Sea K un campo dado, F un subcampo de K y c un elemento de K. Consideremos aquellos elementos de K expresados por polinomios de la forma

$$(1) \qquad f(c)=a_0+a_1c+a_2c^2+\dots+a_nc^n \quad (\text{todo } a_i \text{ en } F).$$

Todo subdominio de K que contenga a F y c contendrá necesariamente a todos estos elementos $f(c)$. Inversamente, el conjunto de tales polinomios es cerrado para la adición, sustracción y multiplicación. Por consiguiente, las expresiones (1) constituyen el subdominio de K engendrado por F y c. Se conviene en designar este subdominio por $F[c]$, con corchetes.

Si $f(c)$ y $g(c)\neq 0$ son expresiones polinómicas análogas a (1), su cociente $f(c)/g(c)$ es un elemento de K, llamado *expresión racional* en c con coeficientes de F. El conjunto de todos estos cocientes es un subcampo; es el campo engendrado por F y c, que convencionalmente se designa por $F(c)$, con paréntesis.

Un campo K se dice *ampliación simple* de su subcampo F si K es engendrado sobre F por un solo elemento c, o sea $K=F(c)$. Los campos $R(\sqrt{2})$, $R(\sqrt[3]{5})$ y $R(\omega)$ discutidos en Cap. II, § 1, son sendos ejemplos de ampliación simple. Puede probarse que toda am-

pliación de cualquier F puede obtenerse por una sucesión finita o transfinita (bien ordenada) de ampliaciones simples.

Fuera del campo de los números racionales, hay algunos números complejos, tales como i, $\sqrt{2}$, $\sqrt[4]{5}$, $\sqrt{-3}$, que satisfacen ecuaciones con coeficientes racionales; pero otros números, tales como π y $e=2,71828...$ no satisfacen a semejantes ecuaciones (excepto la trivial, con todos los coeficientes nulos). Estos últimos números son llamados «transcendentes». Esta importante dicotomía se aplica a los elementos sobre cualquier campo.

DEFINICIÓN. *Sea* K *un campo y* F *un subcampo de* K. *Un elemento* c *de* K *se llamará algebraico sobre* F *si satisface a una ecuación polinómica cuyos coeficientes pertenecen a* F *y no son todos nulos.*

(2) $a_0+a_1c+a_2c^2+\ldots+a_nc^n=0$ *(los* a_i *en* F, *no todos 0).*

Cuando un elemento de K *no sea algebraico sobre* F, *se llamará trascendente sobre* F.

Una ampliación simple $K=F(c)$ se dice algebraica o trascendente sobre F, según que el elemento generador c sea algebraico o trascendente sobre F. La estructura de una ampliación simple trascendente es especialmente fácil de describir.

TEOREMA 1. *Si* c *es trascendente sobre* F, *el subcampo* F(c) *engendrado por* F *y* c *es isomorfo con el campo* F(x) *de todas las formas racionales en una indeterminada* x, *con coeficientes en* F. *El isomorfismo puede ser establecido de tal forma, que* c$\leftrightarrow$x *y* a$\leftrightarrow$a *para todo* a *de* F.

Demostración. Evidentemente, la ampliación $F(c)$ contiene a F y a todas las expresiones $f(c)/g(c)$ con coeficientes en F. Si dos expresiones polinómicas $f_1(c)$ y $f_2(c)$ son iguales en $F(c)$, sus coeficientes deben ser iguales término por término, pues de otro modo la diferencia $f_1(c)-f_2(c)$ nos daría una ecuación polinómica para c de coeficientes no todos nulos, contra la hipótesis de que c es trascendente sobre F. Por consiguiente, la correspondencia $f(c)\leftrightarrow f(x)$ representa el dominio $F[c]$ de manera biunívoca sobre el dominio $F[x]$ de los polinomios en una indeterminada x. Por las reglas para operar con polinomios, esta correspondencia es un isomorfis-

mo. Éste se puede extender, por el Teorema 7 de Cap. II, dando el isomorfismo $f(c)/g(c) \leftrightarrow f(x)/g(x)$ entre $F(c)$ y $F(x)$.

EJERCICIOS

1. Clasificar cada uno de los siguientes complejos como algebraicos o trascendentes sobre el campo R de los números racionales, y decir el porqué: $\sqrt{7}$, $\sqrt[3]{5}$, π^2, $e+3$ (con $e=2{,}71828...$), $i+3$, $e^{2\pi i}$, $\sqrt{2}+i$.
2. Determinar en cada caso de los siguientes, si el número c dado engendra la indicada extensión del campo R de los números racionales; razonar completamente cada respuesta:
   a) $c=1+\sqrt{5}$, en $R(\sqrt{5})$;
   b) $c=1+\sqrt{5}$, en $R(\sqrt[3]{5})$;
   c) $c=\pi+3$, en $R(\pi)$;
   d) $c=\pi^2$, en $R(\pi)$.
3. ¿Qué números en $R(\sqrt{5})$ engendran este mismo campo?
4. a) Si d es un entero no cuadrado perfecto, describir el campo $R(\sqrt{d})$.
   b) Hallar aquellos elementos en $R(\sqrt{d})$ que engendren este mismo campo.
   c) Expresar cada uno de tales elementos como raíces de una ecuación cuadrática con coeficientes en R.

2. Elementos algebraicos sobre un campo

Vamos a investigar la estructura del subcampo de un campo dado K, engendrado sobre F por un elemento algebraico u. Por definición, este elemento debe satisfacer a una ecuación polinómica sobre F, de grado por lo menos igual a 1. El mismo elemento puede satisfacer a muy diferentes ecuaciones; por ejemplo, $\sqrt{2}$ es raíz de $x^2-2=0$, $x^3-2x=0$, $x^4-4=0$, etc. Pero podemos elegir siempre una ecuación «mínima» como sigue:

TEOREMA 2. *Un elemento* u *algebraico sobre un campo* F *es siempre raíz de un polinomio mónico e irreducible en el dominio* F[x] *de todas las formas polinómicas sobre* F. *El polinomio* p(x) *con estas condiciones es único. El elemento* u *será raíz de otro polinomio* g(x) *con coeficientes en* F *si, y sólo si,* g(x) *es múltiplo de* p(x) *en el dominio* F[x].

Demostración. Por definición, el elemento algebraico u es raíz de al menos un polinomio $f(x) \neq 0$ en $F[x]$ (aquí, $f(x) \neq 0$ significa: «los coeficientes de f no son todos nulos»). Si $f(x)$ no es irreducible, admitirá, por el Teor. 12 del Cap. IV, una factorización $f(x)=cp_1(x)...p_m(x)$ en factores mónicos irreducibles, con $c \neq 0$.

Como $f(u)=0$, al menos para un factor p_i resultará $p_i(u)=0$. Se tiene, pues, un polinomio mónico irreducible $p(x)$ con $p(u)=0$. Llamemos n al grado de $p(x)$.

Comprobemos que u no es raíz de ningún polinomio $f(x)=0$ de grado menor que n. Supongamos momentáneamente que un tal polinomio $f(x)$ tenga u como raíz. Ya que $f(x)$ tiene grado inferior al del polinomio irreducible $p(x)$, $f(x)$ y $p(x)$ tendrán como máximo común divisor el 1. Este m. c. d. puede ser expresado (Teor. 10, Capítulo IV) como sigue: $1=t(x)p(x)+s(x)f(x)$, siendo $t(x)$ y $s(x)$ dos polinomios. Sustituyendo aquí x por u resulta $p(u)=0$ y $f(u)=0$, y, por lo tanto, $1=0$, lo que es contradictorio. Por tanto, la igualdad $f(u)=0$ sólo puede ser válida para polinomios $f(x)\neq 0$ cuyo grado sea al menos n, grado de $p(x)$.

Consideremos ahora un polinomio $g(x)$ en $F[x]$ para el cual $g(u)=0$. Efectuando la división de $g(x)$ por $p(x)$, se tiene $g(x)=$ $=q(x)p(x)+r(x)$, donde el resto $r(x)$ tiene grado $n-1$ como máximo. Póngase aquí $u=x$; como $g(u)=0$ y $p(u)=0$, $r(u)$ es también cero. Por la argumentación anterior debe ser $r(x)$ idénticamente cero, y así $g(x)=q(x)p(x)$. Esto demuestra que todo polinomio $g(x)$ que admita la raíz u es idéntico a un múltiplo de $p(x)$. Recíprocamente, es claro que todos los múltiplos de $p(x)$ tienen u por raíz. Como todos estos múltiplos son necesariamente reducibles, el propio $p(x)$ es el único polinomio mónico irreducible con la raíz u. Esto completa la demostración del teorema.

Definición. *El grado* $n=[u:F]$ *de un elemento* u *algebraico sobre un campo* F *es el grado* n *del único polinomio mónico irreducible con coeficientes en* F *que tiene la raíz* u.

Corolario. *Si* u *tiene grado* n *sobre un campo* F, *la igualdad* $a_0+a_1u+\dots+a_{n-1}u^{n-1}=0$ *con coeficientes* a_i *de* F, *se cumplirá si, y sólo si,* $a_0=a_1=\dots=a_{n-1}=0$.

Ahora nos será posible describir el subcampo de K engendrado por F y por uno de sus elementos algebraicos u. Este subcampo $F(u)$ contiene evidentemente el subdominio $F[u]$ de todos los elementos expresables como polinomios $f(u)$ con coeficientes en F [cfr. (1)]. Pero este dominio $F[u]$ resulta ahora un subcampo de K. En efecto, hallaremos un inverso para cada elemento $f(u)\neq 0$ en $F[u]$. El decir que $f(u)\neq 0$ significa que u no es raíz

de $f(x)$ y, por lo tanto, según el Teorema 2, que $f(x)$ no es múltiplo del polinomio irreducible $p(x)$ y, en consecuencia, que $f(x)$ y $p(x)$ son primos entre sí. Por lo tanto, podemos escribir

$$1=t(x)f(x)+s(x)p(x), \tag{8}$$

para convenientes polinomios $t(x)$ y $s(x)$ en $F[x]$. La correspondiente igualdad en $F[u]$ es $1=t(u)f(u)$. Esto prueba que los elementos no nulos $f(u)$ de $F[u]$ tienen un recíproco $t(u)$, el cual es también de forma polinómica (*) en u. Esto prueba que $F[u]$ es un subcampo de K.

Ya que, inversamente, todo subcampo de K que contenga a F y a u contiene con evidencia a todo polinomio $f(u)$ de $F[u]$, vemos que $F[u]$ es el subcampo de K engendrado por F y u. En este subcampo, las reglas para sumar y multiplicar dos polinomios $f(u)$ y $g(u)$ están dadas por el criterio de que las sumas y productos se formen exactamente como para los correspondientes polinomios $f(x)$ y $g(x)$ en una indeterminada x. En otras palabras, la correspondencia $f(x) \to f(u)$ es un homomorfismo. Hemos demostrado así

Teorema 3. *Sea* K *un campo y sea* u *un elemento de* K *algebraico sobre el subcampo* F *de* K. *Entonces, el subcampo* F(u) *engendrado por* F *y* u, *consiste en los elementos de* K *que pueden ser representados como polinomios* f(u) *con coeficientes en* F; *además, la correspondencia* f(x) → f(u) *es un homomorfismo del dominio polinómico* F[x] *al* F(u).

En el caso particular del campo $R(\sqrt{2})$, engendrado sobre el de los racionales por un elemento $\sqrt{2}$ de grado dos, todos los elementos pueden ser escritos como polinomios lineales $a+b\sqrt{2}$. Representaciones similares son posibles en otros campos especiales (Cap. II), de acuerdo con el siguiente resultado general:

Teorema 4. *Si en el Teorema 3, el emento* u *es raíz del polinomio* p(x) *mónico, irreducible sobre* F *(Teorema 2), y de gra-*

(*) Por ejemplo, en $R[\sqrt{3}]$, $1+\sqrt{3}$ tiene el inverso multiplicativo, calculado como sigue, por «racionalización del denominador»:

$$1/(1+\sqrt{3})=(1-\sqrt{3})/(1+\sqrt{3})(1-\sqrt{3})=-(1/2)-(1/2)\sqrt{3}.$$

do n, *entonces, cada elemento del subcampo* F(u) *engendrado por* F *y* u *puede ser representado de modo único como un polinomio*

$$(4) \qquad a_0 + a_1u + \ldots + a_{n-1}u^{n-1}$$

con coeficientes a_i *en* F *y de grado no superior a* n — 1. *Para sumar o restar dos de tales polinomios se suman o restan los correspondientes coeficientes. Para multiplicarlos, se forma el polinomio producto y se calcula el resto de dividir este producto por* p(x).

Demostración. Si $f(u)$ es una expresión polinómica de un elemento del campo $F(u)$, podemos hallar el resto $r(x)$ de la división de $f(x)$ por $p(x)$, siendo $f(x)=q(x)p(x)+r(x)$. Puesto que $p(u)=0$, la sustitución $x=u$ en esta ecuación da $f(u)=r(u)$. Esto prueba que cualquier $f(u)$ puede ser puesto en la forma (4); la unicidad resulta por el Corolario del Teorema 2. Finalmente, las reglas para adición y multiplicación son consecuencia del hecho de ser $f(x) \to f(u)$ un homomorfismo.

Ya que $F(u)$ es una imagen homomorfa del dominio $F[x]$ de formas polinómicas, la teoría del homomorfismo (Cap. XIII) prueba que $F(u)$ está unívocamente determinado, salvo isomorfismos, por el ideal C de todos los elementos de $F[x]$ que tienen por correspondiente al 0. [$f(x) \to f(u)=0$.] En el presente caso, el ideal C es justamente el conjunto de todos los polinomios $f(x)$ que tienen u como raíz, y el anterior Teorema 2 prueba que este conjunto es el formado por todos los múltiplos del polinomio mónico irreducible $p(x)$ anulado por u. El ideal C es así el ideal principal $[p(x)]$ engendrado por $p(x)$, y el campo $F(u)$ puede, por lo tanto, ser descrito como el anillo cociente relativo a este ideal principal en $F[x]$, como sigue:

TEOREMA 5. *Sea* u *un elemento de un campo* K, *algebraico sobre un subcampo* F *y raíz del polinomio mónico* p(x) *irreducible sobre* F *(Teor. 2). Entonces, el subcampo* F(u) *engendrado por* F *y* u *es isomorfo con el anillo cociente* F[x]/(p(x)) *que resulta del anillo de las formas polinómicas con coeficientes en* F *relativamente al módulo ideal de todos los múltiplos de* p(x).

EJERCICIOS

1. Hallar cinco ecuaciones polinómicas distintas para $\sqrt{3}$ y mostrar explícitamente que todas ellas son múltiplos de la ecuación mónica irreducible para $\sqrt{3}$ (sobre el campo R).

2. En la extensión simple $R(u)$ engendrada por una raíz u de $u^3-6u^2+9u+3=0$, expresar cada uno de los siguientes valores mediante los elementos $1, u, u^2$, como en (4): u^4, u^5, $3u^5-u^4+2$, $1/(u+1)$, $1/(u^2-6u+8)$.
3. En la extensión simple $R(u)$ engendrada por una raíz real u de la ecuación irreducible $x^3+2x+2=0$, expresar cada uno de los siguientes valores en la forma (4);

$$(u^2+2)(u^3+3u), \quad u^4(u^4+3u^2+7u+5), \quad 1/u, \quad (u+2)/(u^2+3).$$

4. Representar el campo de los números complejos como un anillo cociente a partir del dominio $R^*[x]$ de todos los polinomios con coeficientes reales.
5. Representar el campo $R(\sqrt{2})$ como un anillo cociente a partir del dominio $R[x]$ de formas polinómicas con coeficientes racionales.
6. Sin utilizar el Teorema 2, demostrar: si u es algebraico sobre F, el polinomio mónico de grado mínimo con raíz u es irreducible sobre F.
7. Sin utilizar el Teorema 2, probar directamente: si u es un elemento de un campo K, y F un subcampo de K, entonces el conjunto de todos los polinomios $g(x)$ con coeficientes en F, que tienen u como raíz, es un *ideal* de $F[x]$.

3. Adjunción de raíces

La ampliación de un campo dado F puede ser estudiada desde dos puntos de vista. Se puede partir de un subcampo F de un campo más amplio K, y estudiar el campo $F(u)$ engendrado, sin salir de K, por F y una raíz u en K de un polinomio con coeficientes de F. Este es el camino *concreto* adoptado en lo precedente. O bien, con método distinto, se puede partir simplemente del campo F y de una ecuación polinómica $p(x)=0$ con coeficientes de F, e intentar ampliar F a un campo más extenso en el que resulten incluídas las raíces de esta ecuación. Este es el camino *abstracto*, empleado en el Cap. V cuando se construyó el campo complejo C, partiendo del campo real R^*, por adjunción de una raíz «imaginaria» del polinomio $f(z)=z^2+1$.

El método abstracto también puede ser ilustrado por la construcción de un campo finito. Partamos del campo J_3 de las tres clases residuales de los enteros 0, 1 y 2, módulo 3; procuremos agregar un nuevo símbolo u con el cual se engendre una ampliación K en la que se satisfaga la ecuación $u^2=u+1$. La ampliación K, si existe, debe consistir, por el Teorema 4, en los nueve elementos $a+bu$, donde los elementos a y b son escogidos entre las clases residuales 0, 1 y 2. Teniendo estos nueve entes $a+bu$, inten-

temos hacerlos pertenecer a un campo. La suma de dos de ellos estará dada por la regla

$$(a+bu)+(c+du)=(a+c)+(b+d)u.$$

Para calcular el producto de dos de estos elementos, los «multiplicaremos» del modo natural y utilizaremos la ecuación $u^2=u+1$. El resultado

$$(a+bu)(c+du)=(ac+bd)+(ad+bc+bd)u,$$

es siempre uno de los nueve elementos dados. Y es posible comprobar detalladamente que estos nueve entes, bajo estas dos operaciones, satisfacen a todos los postulados para un campo, de acuerdo con lo que afirma el Teor. 6 siguiente. En particular, las inversas de los elementos no nulos son las siguientes:

1	2	u	$2u$	$1+u$	$1+2u$	$2+u$	$2+2u$
1	2	$2+u$	$1+2u$	$2+2u$	$2u$	u	$1+u$

Por esta construcción hemos obtenido el campo $J_3(u)$ engendrado por u sobre el campo J_3 de las clases residuales. Este es un ejemplo sencillo de campo finito (ver Cap. XV, § 6).

En general, se pueden adjuntar a su campo F raíces «ficticias» de cualquier polinomio $f(x)$. Si $f(x)$ no es irreducible, poseerá un factor mónico irreducible $p(x)$; también cada raíz u de $p(x)=0$ satisface a $f(x)=0$. Por lo tanto, es suficiente construir las extensiones abstractas de F por adjunción de raíces de una ecuación irreducible. Vamos a demostrar que ello es posible.

Teorema 6. *Para todo polinomio* p(x) *irreducible sobre un campo* F, *existe un campo tal, que es una extensión algebraica simple de* F *engendrada por una raíz* u *de* p(x).

Demostración. El Teorema 5 asegura que si existe una tal extensión simple $F(u)$, deberá ser isomorfa con el anillo cociente $F[x]/(p(x))$. Inversamente, comencemos nuestra construcción con el dominio $F[t]$ de todas las formas polinómicas en una indeterminada t (nueva) y con el ideal (principal) $C=(p(t))$ de todos los múltiplos de $p(t)$ en este dominio. Cada polinomio $g(t)$ determi-

na una clase $(g(t))' = g(t) + C$, y estas clases forman un anillo $F(t)/(p(t))$, el cual es imagen homomorfa de $F[t]$ bajo la correspondencia $g(t) \to (g(t))'$. El único elemento de F que se corresponde con el ideal $(p(t))$ es cero; así, este homomorfismo actúa sobre los elementos del campo base F como un isomorfismo $a \leftrightarrow a'$ (cf. Capítulo XIII, Teorema 3). Identificaremos cada elemento a de F con su correspondiente clase a' en el anillo-cociente. Cualquier clase

$$(g(t))' = (a_0 + a_1 t + \ldots + a_m t^m)' = a_0 + a_1 t' + \ldots + a_m t'^m$$

en el anillo-cociente, es entonces un polinomio en la clase t', con coeficientes en F, luego el anillo-cociente puede ser engendrado como $F[t']$. Puesto que $p(t)$ está en la clase 0, $p(t') = 0$, y el generador t' satisface a la ecuación polinómica irreducible dada. Falta sólo probar que $F[t']$ es un campo. Por el Teorema 6, Cor. 1, del Capítulo XIII, $F[t']$ es, desde luego, un anillo conmutativo con elemento unidad.

Por lo dicho, basta hallar un inverso para cada $g(t') \neq 0$. Pero $g(t') \neq 0$ significa que $g(t)$ no es divisible por el irreducible $p(t)$, por lo tanto, que $g(t)$ y $p(t)$ son primos relativos y, por lo tanto, que existen en $F[t]$ polinomios $r(t)$ y $s(t)$ tales, que

$$1 = r(t)g(t) + s(t)p(t).$$

Aplicando el homomorfismo, resulta que $1 = r(t')g(t')$, obteniendo así el inverso que se pretendía, y la demostración está completada.

Este proceso de adjunción puede ser aplicado con cualquier campo F como base. Si F es el campo R^* de los números reales, y $p(x)$ el polinomio $x^2 + 1$ irreducible sobre R^*, entonces, la construcción da un campo $R^*(t')$ engendrado por una cantidad t' con $t'^2 = -1$. Esta cantidad t' actúa como $i = \sqrt{-1}$, y el campo $R^*(t')$ es isomorfo con el de los números complejos. Tenemos así una ligera variante de la construcción estudiada en el Cap. V, de los números complejos partiendo de los números reales.

Si F es el campo J_p de los enteros módulo p, y si $p(x)$ es un polinomio irreducible sobre F, la construcción anterior dará un campo que consistirá en los elementos $a_0 + a_1 t' + \ldots + a_{n-1} t'^{n-1}$. Como sólo pueden escogerse p valores distintos para cada coeficiente a_i, resulta que el campo construído es finito y consta de p^n elementos.

Si $F = C(x)$ es el campo de todas las formas racionales en una indeterminada x, con coeficiente en el campo C de los números

complejos, el polinomio $p(t)=t^2-(x^2-1)(x^2-4)$ es irreducible (como polinomio en t) sobre $C(x)$. La construcción del Teorema 6 produce un campo $F(y)=C(x, y)$ engendrado por una raíz $y=\sqrt{(x^2-1)(x^2-4)}$ del polinomio dado $p(t)$. Este campo consiste en los elementos de la forma

$$r(x)+s(x)y=r(x)+s(x)\sqrt{(x^2-1)(x^2-4)},$$

donde $r(x)$ y $s(x)$ son formas racionales en x. Estos elementos son llamados funciones algebraicas de x, pues se expresan en función de x por operaciones algebraicas. Cualquier extensión algebraica simple del campo $C(x)$ constituye un *campo de funciones algebraicas*. En particular, el campo $C(x, y)$ es un *campo de funciones elípticas*, porque es engendrado por la cantidad y que aparece en la integral elíptica

$$\sqrt{(x^2-1)(x^2-4)}dx$$

Si la construcción del Teorema 6 se aplica a un polinomio ordinario, tal como x^3-5, irreducible sobre el campo R de los racionales, resulta un cierto campo abstracto $R(u)$ engendrado por una raíz u de $x^3-5=0$. Este campo es conceptualmente diferente del $R(\sqrt[3]{5})$ engendrado por la raíz positiva de la misma ecuación. Sin embargo, los dos campos $R(u)$ y $R(\sqrt[3]{5})$ son algebraicamente indistinguibles, a causa de ser isomorfos. Esto es una consecuencia del siguiente teorema general.

TEOREMA 7. *Si los campos* F(u) *y* F(w) *son ampliaciones algebraicas simples del mismo campo* F, *engendrados respectivamente por las raíces* u *y* w *del mismo polinomio* p(x) *irreducible sobre* F, *entonces* F(u) *y* F(w) *son isomorfos. Precisamente, se trata de un isomorfismo de* F(u) *a* F(w) *en el cual* u *se corresponde con* w *y cada elemento de* F *se corresponde consigo mismo.*

Demostración. Vamos a establecer una correspondencia biunívoca entre $F(u)$ y $F(w)$, la cual conserve sumas y productos. Los elementos de $F(w)$, semejantemente a los de $F(u)$, pueden ser representados como polinomios $f(w)=a_0+a_1w+\ldots+a_kw^k$, con coeficientes de F. Esto sugiere la correspondencia $f(u)\leftrightarrow f(w)$, en la cual cada polinomio en u se corresponde con el polinomio en w que tiene los mismos coeficientes. De este modo, a las sumas y produc-

tos de polinomios les corresponden las sumas y productos de los homólogos. Falta sólo verificar que la correspondencia es biunívoca, o que $f(u)=g(u)$ si, y sólo si, $f(w)=g(w)$. Pero $f(u)=g(u)$ significa que $f(u)-g(u)=0$, lo cual significa, según el Teorema 2, que $p(x) \mid [f(x)-g(x)]$, lo cual a su vez significa que $f(w)=g(w)$, ya que w satisface también a $p(x)=0$. El razonamiento es también válido en sentido contrario (permutándose u con w), y la conclusión queda establecida.

Este isomorfismo significa, sencillamente hablando, que dos raíces cualesquiera de un polinomio irreducible $p(x)$ tienen el mismo comportamiento algebraico, y que todas las propiedades algebraicas de una raíz u pueden ser deducidas de la ecuación irreducible a que satisface. Existen muchos ejemplos de este isomorfismo. Por ejemplo, el campo $C=R^*(i)$ de los números complejos es engendrado sobre el R^* de los números reales por cualquiera de las dos raíces $\pm i$ de la ecuación $x^2+1=0$. Por lo tanto, existe, según el Teorema 7, un automorfismo de C que transporta i sobre $-i$. Este automorfismo $a+bi \leftrightarrow a-bi$ es precisamente la correspondencia entre un número complejo y el complejo conjugado.

EJERCICIOS

1. Mostrar un automorfismo de cada uno de los siguientes campos: $R(\sqrt{2})$, $R(\sqrt{-3})$, $R(i)$.
2. Mostrar un campo de números complejos isomorfo con cada uno de los campos $R(\sqrt[3]{5})$, $R(\sqrt[4]{2})$.
3. Demostrar que x^3+x-1 es irreducible sobre el campo J_5 de los enteros módulo 5. Si una raíz de este polinomio se adjunta a J_5, ¿cuántos elementos tendrá el campo obtenido?
4. a) Formar polinomios de grados 2 y 3 irreducibles sobre el campo J_2 de los enteros módulo 2.
   b) Construir las tablas para la adición y multiplicación en un campo con cuatro elementos.
5. a) Mostrar que el campo de nueve elementos construido en el texto tiene característica 3.
   b) Mostrar explícitamente el isomorfismo $a \leftrightarrow a^3$ en dicho campo (Capítulo XIII, Teor. 15).
6. a) Hallar todos los polinomios cuadráticos irreducibles sobre el campo J_3.
   b) Demostrar que, excepto los isomorfos con él, hay un solo campo con nueve elementos. (*Sugerencia:* Demostrar que cualquier elemento de un campo tal, satisface a una ecuación cuadrática sobre el subcampo J_3.)

7. Demostrar que el polinomio en t, $t^4-(x^2-1)(x^2-4)$, es irreducible sobre el campo $C(x)$. (*Sugerencia:* Utilizar los resultados de § 8, Cap. IV.)
8. Probar que el campo de funciones elípticas $C(x, y)$ citado en el texto. puede ser engendrado sobre $C(x)$ por una raíz z de la ecuación $t^2=$ $=(x^2-4)/(x^2-1)$.
9. Si $g(t)$ es un polinomio reducible, ¿cuáles son los elementos del anillo cociente $F[t]/[g(t)]$ que tienen inverso?
10. Mediante los resultados del Capítulo XIII, dar otra demostración de que $F[t]/[p(t)]$ es un campo.

4. Ampliaciones finitas. Grado

En una ampliación simple $F(u)$ engendrada por un elemento u de grado n, cada elemento w tiene una representación única en la forma

$$w=a_0+a_1u+\ldots+a_{n-1}u^{n-1}, \tag{5}$$

con coeficientes en F. Esta única representación se asemeja, desde luego, a la representación de un vector mediante los vectores de una «base» $1, u, \ldots, u^{n-1}$. Además, el corolario del anterior Teorema 2 es muy semejante a la condición para la independencia lineal de estos elementos. Todo esto sugiere una aplicación del concepto de espacio vectorial.

En general, cualquier extensión K de un campo F puede ser considerada como un *espacio vectorial* sobre F: es desconocida la multiplicación de los elementos de K, y como operaciones del espacio vectorial se emplearán la adición de dos elementos de K y la multiplicación «escalar», esto es, multiplicación de los elementos de K por los elementos de F. Todos los postulados característicos del espacio vectorial son satisfechos por esta adición y multiplicación escalar. Si este espacio vectorial K tiene dimensión finita, entonces el campo K se llama *ampliación* (o *extensión*) *finita de* F, y la dimensión n del espacio vectorial será llamada el *grado* $n=[K:F]$ de la extensión.

Por ejemplo, el campo complejo $C=R^*(i)$ es un espacio vectorial de dos dimensiones sobre el subcampo real R^* (recordar el diagrama de Argand). El campo $R(\sqrt[3]{5})$ engendrado por los números racionales y una raíz cúbica de 5 es un espacio vectorial tridimensional sobre el subcampo racional R, etc. En general, el resultado del Teorema 4 sobre ampliaciones algebraicas simples puede ser nuevamente enunciado, refiriéndonos a la dimensión, como sigue:

TEOREMA 8. *El grado* n *de un elemento algebraico* u *sobre un campo* F *es igual a la dimensión de la ampliación* F(u) *considerada como un espacio vectorial sobre* F. *Este espacio vectorial tiene una base* 1, u, u^2, ..., u^{n-1}.

En § 5 veremos de qué modo el espacio vectorial puede ser empleado para analizar las ampliaciones de un campo F obtenidas por adjunción de varios elementos algebraicos distintos. Pero antes de discutir tales ampliaciones «múltiples», veremos como el espacio vectorial facilita el comparar las ecuaciones irreducibles sobre F satisfechas por diferentes elementos de la misma extensión algebraica simple $F(u)$.

Un hecho fundamental relativo a los espacios vectoriales es la invariancia de la dimensión (dos bases cualesquiera del espacio tienen el mismo número de elementos). Este hecho puede ser aplicado al caso especial de la ampliación finita de un campo, como sigue:

COROLARIO. *Si dos elementos algebraicos* u *y* v *sobre un campo* F *engendran la misma ampliación* F(u)=F(v), *ambos elementos tienen el mismo grado sobre* F.

Una ampliación algebraica es siempre finita, e inversamente, cualquier extensión finita consta de elementos algebraicos.

TEOREMA 9. *Cualquier elemento* w *de un campo* K *ampliación finita de* F *es algebraico sobre* F *y satisface a una ecuación irreducible sobre* F *de grado* n *a lo más, donde* n=[K : F] *es el grado de la ampliación dada.*

Demostración. Las $n+1$ potencias 1, w, w^2, ..., w^n del elemento dado w son elementos del espacio vectorial n-dimensional K, luego deben ser linealmente dependientes sobre F (Cap. VII, Teorema 7, Corolario 2). Por lo tanto, existe una relación lineal $b_0+b_1w+\dots+b_nw^n=0$, con algún coeficiente no nulo. Interpretada como un polinomio, esta relación implica que w es algebraico sobre F, como debíamos demostrar.

COROLARIO. *Cualquier elemento de una ampliación algebraica simple* F(u). *es algebraico sobre* F.

Esta importante conclusión nos asegura que en una ampliación algebraica no puede aparecer nunca un elemento trascendente.

Trabajando con una ampliación algebraica simple particular $F(u)$, el polinomio irreducible $p(x)$ correspondiente a u puede utilizarse sistemáticamente porque, por el Teor. 2, un elemento $g(u)$ de la ampliación es cero si, y sólo si, el polinomio $g(x)$ es divisible por $p(x)$. Supongamos, por ejemplo, que $R(u)$ es una ampliación de grado 3 sobre el campo R de los números racionales, engendrada por una raíz u de x^3-2x+2. Este polinomio es irreducible por el criterio de Eisenstein (Cap. IV, § 9). El elemento $w=u^2-u$ en esta ampliación $R(u)$ debe satisfacer a alguna ecuación de grado 3 a lo más. Para hallar esta ecuación expresaremos las potencias $w^2=u^4-2u^3+u^2$ y $w^3=u^6-3u^5+3u^4-u^3$ como sendas funciones lineales de 1, u y u^2, según el Teor. 4. Esto se logra por aplicación repetida de la ecuación dada $u^3=2u-2$, obteniéndose

$$w=u^2-u, \qquad w^2=3u^2-6u+4, \qquad w^3=16u^2-28u+18.$$

Para obtener la relación lineal que debe existir entre 1, w, w^2 y w^3, resolveremos las dos primeras ecuaciones, lineales en w y w^2, resultando

(6) $$u=-w^2/3+w+4/3, \qquad u^2=-w^2/3+2w+4/3.$$

Estos valores, sustituídos en la expresión de w^3, dan la ecuación deseada

$$w^3-4w^2-4w-2=0.$$

Esta ecuación es irreducible sobre R, por el teorema de Eisenstein. Inversamente, las ecuaciones (6) demuestran que u está en $R(w)$, así que $R(u)=R(w)$; u y w engendran, pues, la misma ampliación y, por el corolario al Teorema 8, tienen el mismo grado 3 sobre R. Esto significa que cualquier ecuación de grado 3 para w debe ser irreducible.

EJERCICIOS

1. Cada uno de los siguientes números pertenece a una ampliación algebraica simple sobre R. Hallar en cada caso la ecuación mónica irreducible satisfecha por el número. a) $2+\sqrt{3}$; b) $\sqrt[3]{5}+\sqrt{5}$; c) $\sqrt[3]{2}+\sqrt[3]{4}$; d) u^2-1, donde u satisface a $u^3=2u+2$; e) u^2+u, donde $u^3=-3u^2+3$.
2. Probar que cualquier ampliación finita del campo R^* de los números reales, o bien coincide con R^*, o bien es isomorfa con el campo C de los números complejos.
3. Probar que el campo de todos los números complejos no tiene ampliaciones finitas.

4. a) Si K es una ampliación de grado 2 sobre el campo R de los racionales, demostrar que $K=R(\sqrt{d})$, donde d es un entero no cuadrado y sin divisores enteros cuadrados perfectos.
   b) ¿Qué subsiste de este teorema, si R se reemplaza por un campo F de característica infinita? ¿Y si lo es por un campo de característica cualquiera?
5. ¿Es el campo $F(x)$ de formas racionales en la indeterminada x una ampliación finita de F? ¿Por qué?
6. Demostrar que el número de elementos de un campo finito de característica p es una potencia de p.
7. a) Demostrar que hay exactamente $(p^2-p)/2$ polinomios cuadráticos mónicos irreducibles sobre el campo J_p de los enteros módulo p.
   b) Demostrar que para cada p existe un campo de característica p con p^2 elementos.
*8. Demostrar que hay exactamente $(p^3-p)/3$ polinomios cúbicos mónicos irreducibles sobre el campo J_p de los enteros módulo p.
*9. Sea F cualquier campo contenido en un dominio de integridad D. Demostrar:
   a) D es un espacio vectorial sobre F;
   b) Si, como espacio vectorial, D tiene dimensión finita sobre F, D es un campo.

5. Extensiones algebraicas reiteradas

Pueden construirse ampliaciones finitas de un campo F mediante sucesivas ampliaciones simples. Si F tiene característica ∞, puede probarse que cualquiera de estas ampliaciones reiteradas puede engendrarse por un solo elemento convenientemente elegido. Omitiremos esta demostración (*) y discutiremos directamente las propiedades de la ampliación reiterada. En general, si K es una ampliación de F que contiene a los elementos $c_1, c_2, \ldots, c_r$, el símbolo $F(c_1, \ldots, c_r)$ denota el subcampo de K engendrado por $c_1, \ldots, c_r$ y los elementos de F (el subcampo consta de todos los elementos racionalmente expresados sobre F mediante $c_1, \ldots, c_r$). Así, $F(c_1, c_2)$ es la ampliación simple $L(c_2)$ de la ampliación simple. $L=F(c_1)$.

Las ampliaciones algebraicas reiteradas aparecen en la solución de ecuaciones, donde frecuentemente es útil introducir apropiadas ecuaciones auxiliares. Por ejemplo, la ecuación $x^4-2x^2+9=0$ puede escribirse

$$x^4-2x^2+9=(x^4-6x^2+9)+4x^2=(x^2-3)^2+4x^2=0.$$

(*) Una demostración se encuentra en Cap. VII de L. Weisner, *Introduction to the Theory of Equations*, New York, 1938. Se recomiendan también los problemas de tal capítulo.

La ecuación, por lo tanto, es $[(x^2-3)/2x]^2=-1$. Esta fórmula indica que cualquier campo que contenga la raíz u de la ecuación dada, contiene también la raíz $i=(u^2-3)/2u$ de la ecuación $y^2=-1$. Si adjuntamos la cantidad auxiliar i al campo racional R, la ecuación original resulta reducible sobre $R(i)$,

$$x^4-2x^2+9=(x^2-3+2xi)(x^2-3-2xi).$$

La fórmula usual muestra que el factor $x^2-3-2ix$ tiene como raíz $u=i+\sqrt{2}$. La ecuación propuesta, por lo tanto, tiene una raíz en el campo $K=R(i, \sqrt{2})$. El campo K es obtenido por adjunción a R, primero de $\sqrt{2}$, y después de i. El campo intermedio $R(\sqrt{2})$ consta de números reales, luego no puede contener a i. La ecuación cuadrática $y^2+1=0$ que da i debe, por lo tanto, permanecer irreducible sobre el campo real $R(\sqrt{2})$, así que la extensión $R(\sqrt{2}, i)$ tiene grado 2 sobre $R(\sqrt{2})$, y una base son los elementos 1 e i. El campo $R(\sqrt{2})$, a su vez, tiene una base 1, $\sqrt{2}$ sobre R. Por consiguiente, cualquier elemento w en el campo total $R(\sqrt{2}, i)$ puede ser expresado como

$$w=(a+b\sqrt{2})+(c+d\sqrt{2})i=a+b\sqrt{2}+ci+d\sqrt{2}i, \tag{7}$$

con coeficientes racionales a, b, c y d. Los cuatro elementos 1, $\sqrt{2}$, i, $\sqrt{2}i$ aparecen así formando una base para la ampliación total $R(\sqrt{2}, i)$ sobre R. Este método de composición de bases puede ser establecido en general como sigue:

TEOREMA 10. *Si los elementos* $u_1, \ldots, u_n$ *forman una base de una extensión finita* K *de* F, *y asimismo los* $w_1, \ldots, w_m$ *constituyen una base de una ampliación* L *de* K, *entonces los* mn *productos* $u_i w_j$ $(i=1, \ldots, n;\ j=1, \ldots, m)$ *forman una base de* L *sobre* F.

Demostración. Cualquier elemento y en L puede ser representado como una combinación lineal $y=\sum_j r_j w_j$ de la base dada, con coeficientes r_j en K. Cada coeficiente r_j es a su vez una combinación $r_j=\sum_i a_{ij}u_i$ de los elementos base de K, con todos los a_{ij} en F. Por sustitución de estos valores, $y=\sum_j\sum_i a_{ij}u_i w_j$ aparece como una combinación lineal de los anunciados elementos bases $u_i w_j$, con coeficientes en F. El mismo razonamiento prueba que estos mn ele-

mentos son linealmente independientes sobre F, luego constituyen una base de K.

Muchas consecuencias se deducen del Teorema 10. En primer lugar, se puede establecer el resultado sin referencia a la base particular empleada, como sigue:

COROLARIO 1. *Si* K *es una ampliación finita de* F, *y* L *una ampliación finita de* K, *entonces* L *es una ampliación finita de* K *y su grado es*

(8) $$[L:F]=[L:K][K:F] \quad (L \geqslant K \geqslant F).$$

COROLARIO 2. *Si* K *es una ampliación finita de grado* n$=$[K : F] *sobre* F, *cualquier elemento* u *de* K *tiene sobre* F *un grado que es divisor de* n.

Demostración. El elemento u engendra una ampliación simple $F(u)$; por lo tanto, por (8), $n=[K:F(u)][F(u):F]$, donde el segundo factor es el grado de u que dice el enunciado.

COROLARIO 3. *Un elemento* u *de una ampliación finita* K$\geqslant$F *engendra toda la ampliación si, y sólo si,* [K : F]$=$[u : F].

Demostración. Si u satisface sobre F una ecuación irreducible de grado $[K:F]$, entonces u engendra un subcampo $F(u)$ de grado n sobre F. Por (8), este subcampo debe incluir a todo K.

COROLARIO 4. *Si* K$=$F(y_1, y_2, ..., y_r) *es un campo engendrado por* r *cantidades* y_i, *donde cada* y_i *es algebraica sobre el campo* F(y_1, ..., y_{i-1}) *engendrado por las* i$-$1 *cantidades precedentes, entonces* K *es una extensión finita de* F *y cada elemento de* K *es algebraico sobre* F.

Demostración. Todos los grados $[F(y_1, \ldots, y_{i-1}, y_i):F(y_1, \ldots, y_{i-1})]$ son finitos; luego, por el Corolario 1, el grado $[K:F]$ es finito. Por el Teorema 9, cualquier elemento en K será algebraico sobre F.

COROLARIO 5. *Si* p(x) *es un polinomio cúbico irreducible sobre un campo* F *y si* K *es una ampliación de* F *de grado* 2^m, *el polinomio* p(x) *es irreducible sobre* K.

Este corolario significa, en particular, que una ecuación cúbica irreducible no puede ser resuelta por el cálculo sucesivo de raíces

cuadradas, pues la adjunción de una raíz cuadrada al campo F, o bien no producirá ninguna ampliación, o bien será una extensión de grado 2, así que el campo $K=F(\sqrt{a}, \sqrt{b}, \sqrt{c}, \ldots)$ que se obtenga por cualquier número de raíces cuadradas tiene por grado alguna potencia 2^m de 2. Por el Corolario 5, esta ampliación no contendrá ninguna raíz de la ecuación cúbica irreducible.

Para demostrar ahora dicho Corolario 4, supongamos $p(x)$ reducible sobre el campo K de grado 2^m. Entonces el polinomio de tercer grado $p(x)$ debe tener al menos un factor lineal $x-u$, así que K contiene una raíz u de $p(x)$. Pero un tal elemento u de grado 3 sobre F, no puede estar contenido en un campo K de grado 2^m sobre F, por el Corolario 2. Esto prueba que $p(x)$ es irreducible.

Este corolario es el fundamento algebraico del teorema que afirma la imposibilidad de resolver el clásico problema de la duplicación del cubo, o el de la trisección del ángulo, con sólo regla y compás. En efecto: una construcción geométrica puede traducirse a términos analíticos. Los datos del problema consisten en cierto número de puntos y de líneas; relativamente a ciertos ejes cartesianos, las coordenadas de estos puntos (y las razones entre los coeficientes de las ecuaciones de estas líneas) son un conjunto de números reales, los cuales engendran un cierto campo F de números reales. Cada paso de una construcción con regla y compás proporciona nuevos puntos y líneas. Puede demostrarse (*) que el nuevo campo de números que les corresponde es el mismo F o es una ampliación cuadrática de F. Por lo tanto, las construcciones repetidas proporcionan un conjunto de puntos y de líneas al cual corresponde un campo K de grado 2^m sobre F.

Consideremos ahora la duplicación del cubo. Los datos son un par de ejes coordenados y un segmento unidad sobre uno de estos dos ejes, con un cubo que tiene por arista este segmento. El problema es construir un cubo de volumen doble. El lado de este cubo verificará la ecuación $x^3-2=0$. Por el teorema de Eisenstein, esta ecuación es irreducible sobre el campo R de los racionales (que es el campo asociado a los datos). En cualquier campo K correspondiente a cualesquiera construcciones con regla y compás, el poli-

(*) *Esto resulta considerando que la ecuación del círculo (compás) es cuadrática y la de la recta (regla) es lineal. Para la discusión detallada nos remitimos a los textos de Teoría de Ecuaciones, como, por ejemplo, Weisner, loc. cit., o L. E. Dickson, New First Course in the Theory of Equations*, New York, 1939.

nomio x^3-2 permanecerá irreducible, por Corolario 1; luego es imposible construir un segmento del eje x cuya longitud determine el cubo doble.

El problema de la trisección se trata de manera análoga; lo esencial consiste en escribir la ecuación trigonométrica para el coseno del tercio de un ángulo en función del coseno del ángulo entero. Salvo para ángulos especiales, resulta así una ecuación cúbica irreducible.

EJERCICIOS

1. En el Teorema 10, demostrar detalladamente que los mn elementos $u_i w_j$ son independientes sobre F.
2. Demostrar que el polinomio x^4-2x^2+9 considerado en el texto, es irreducible sobre R. [*Sugerencia:* Atender al grado de $R(\sqrt{2}, i)$.]
3. Si $p(x)$ es un polinomio de grado q y es irreducible sobre F, y si K es una ampliación finita de F de grado primo con q, demostrar que $p(x)$ es irreducible sobre K.
4. Determinar el grado de cada una de las siguientes ampliaciones múltiples del campo R de los números racionales:
   a) $R(\sqrt{3}, i)$;
   b) $R(\sqrt[3]{5}, \sqrt{-2})$;
   c) $R(\sqrt{18}, \sqrt[4]{2})$;
   d) $R(\sqrt{8}, 3+\sqrt{50})$;
   e) $R(\sqrt[4]{2}, u)$, con $u^4+6u+2=0$;
   f) $R(\sqrt{3}, \sqrt{-5}, \sqrt{7})$;
   g) $R(\sqrt{3}, \sqrt{2})$.
5. Dar una base sobre R para cada campo del ejercicio anterior.
6. Determinar razonadamente cuáles de los polinomios que siguen son irreducibles sobre el campo que se indica:
   a) x^3+3, sobre $R(\sqrt{7})$;
   b) x^2+1, sobre $R(\sqrt{-2})$;
   c) x^3+8x-2, sobre $R(\sqrt{2})$;
   d) x^3+3x^2-9x-6, sobre $R(\sqrt{7}, \sqrt{5}, 1+i)$.
7. Determinar en los casos que siguen si el número u dado engendra la ampliación que se indica del campo R de los números racionales, desarrollando la correspondiente prueba:
   a) $u=\sqrt[3]{7}$, en $R(\sqrt[3]{7})$;
   b) $u=\sqrt{2}+\sqrt{5}$, en $R(\sqrt{2}, \sqrt{5})$;
   c) $u=2+\sqrt[3]{9}$, en $R(\sqrt[3]{3})$;
   d) $u=\sqrt{2}-1/(1+\sqrt{2})$, en $R(\sqrt{2})$;
   e) $u=v^2+v+1$, en $R(v)$, con $v^3+5v-5=0$.
8. ¿Es trascendente o algebraico sobre R el número $c=\pi^4+5\pi^2+2\pi-14$? ¿Por qué?

9. Si K es una ampliación de F de grado primo, demostrar que todo elemento que esté en K pero no en F, engendra todo K sobre F.
10. a) Plantear la ecuación cúbica que da $\cos\theta$ en función de $\cos 3\theta$.
    b) Demostrar que esta ecuación es irreducible cuando $3\theta=60°$ (ello significa que el ángulo de 60° no puede ser trisecado con regla y compás).

6. Números algebraicos

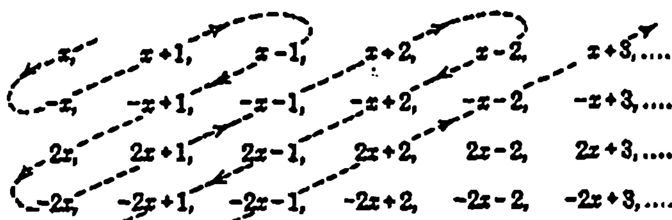
Un *número algebraico* u es un número complejo que satisface a una ecuación polinómica con coeficientes racionales no todos nulos:

$$(9)\qquad a_0+a_1u+a_2u^2+\ldots+a_nu^n=0 \quad (\text{todo } a_i \text{ en } R;\ \text{algún } a_i\neq 0).$$

Dicho de otro modo, un número algebraico es un número complejo que es algebraico sobre el campo R de los racionales. Discutiendo la extensión de campos se han utilizado repetidos ejemplos de números algebraicos, tales como i, $\sqrt{-2}$, $\sqrt[5]{3}$ y ω.

TEOREMA 11. *El conjunto de todos los números algebraicos es numerable.*

La demostración de este aserto requiere que describamos un método que permita la enumeración (o el alistamiento, por así decirlo) de los números algebraicos. Primeramente haremos una lista con todas las ecuaciones a que ellos satisfacen. Observemos antes que la ecuación (9) puede ser multiplicada por el m. c. m. de los denominadores de los coeficientes, resultando una ecuación con coeficientes enteros no todos nulos, de los cuales el primero puede siempre suponerse positivo. Sabemos ya que los coeficientes enteros y positivos de estos polinomios pueden ser enumerados, como en esta lista: 0, +1, —1, +2, —2, +3, —3, ... Todos los polinomios lineales con coeficientes enteros pueden también ordenarse en un rectángulo, como se ve a continuación:



Tomando cada elemento del cuadro en el orden que indican las diagonales, dispondremos de una lista que incluye todos los polinomios lineales. Estos quedan, pues, ordenados en la sucesión siguiente:

$$x, -x, x+1, x-1, -x+1, 2x, -2x, 2x+1, -x-1, \ldots$$

hora podremos disponer una ordenación rectangular de todos los polinomios de segundo grado, sin más que agregar los varios términos mx^2 a los diversos términos de la anterior lista. Y de este rectángulo, por el mismo proceso, se obtiene una fila en la que aparecerán todos los polinomios de segundo grado; y de aquí se pasará a los de grado más elevado. Ahora podremos disponer una sobre otra las listas que así hemos obtenido, de modo que los elementos de la fila n-ésima serán los polinomios de grado n. Desarrollando este rectángulo según las diagonales, se obtiene, como antes, una sola alineación con todos los polinomios. En esta alineación reemplacemos cada polinomio por sus raíces y suprimamos las que aparezcan duplicadas. El resultado es una lista en que figuran todas las raíces de polinomios con coeficientes enteros; ésta es la deseada enumeración de todos los números algebraicos.

Una consecuencia es, que el conjunto de los números algebraicos reales es numerable. Pero el proceso diagonal de Cantor demuestra que el conjunto de todos los números reales no es numerable (Cap. XII, Teor. 5). Luego este conjunto debe ser más amplio que el de los números algebraicos. Esto demuestra de modo indirecto la existencia de números reales transcendentes. El resultado se enuncia como sigue:

Corolario. *No todos los números reales son algebraicos.*

La demostración de Cantor fué primeramente rechazada por muchos matemáticos, a causa de que no permite construir efectivamente un número transcendente. En la actualidad es más generalmente aceptada, pero es posible dar otras demostraciones más explícitas de este Corolario (ver los siguientes Ejercicios 10-13).

Teorema 12. *El conjunto de todos los números algebraicos es un campo.*

Demostración. Solamente necesitamos probar que la suma, producto, diferencia y cociente de dos números algebraicos u y

$v \neq 0$, son también números algebraicos. Pero todas estas combinaciones están contenidas en el subcampo $R(u, v)$ del campo de los números complejos engendrados por u y v. Como u es algebraico sobre R, $R(u)$ es una ampliación finita de R; como v es algebraico sobre $R(u)$, $R(u, v)$ es finita sobre $R(u)$. Luego, por el Teorema 10, $R(u, v)$ es finita sobre R, así que cada uno de sus elementos es un número algebraico (Teorema 9).

Un campo F se llama algebraicamente cerrado (*) (o algebraicamente completo) si cualquier polinomio con coeficientes en F tiene sus raíces en F. En tal caso, cualquier polinomio $f(x)$ sobre F tiene una raíz c en este campo, y por ende, un factor lineal $x - c$. En consecuencia, los únicos polinomios irreducibles sobre F son los lineales, y cualquier polinomio sobre un campo algebraicamente cerrado puede ser descompuesto en un producto de factores lineales (como en la fórmula (11) del Cap. V). Además, no existe ninguna ampliación algebraica de F, excepto el propio F. Se concluye que un campo es algebraicamente cerrado si, y sólo si, no tiene ampliaciones algebraicas propiamente tales. El teorema fundamental del Álgebra (Cap. V, Teor. 5) afirma que el campo de los números complejos es algebraicamente cerrado.

TEOREMA 13. *El campo A de los números algebraicos es algebraicamente cerrado.*

Demostración. Consideremos una ecuación polinómica $x^n + u_{n-1}x^{n-1} + \ldots + u_0 = 0$ cuyos coeficientes sean números algebraicos u_i contenidos en A. Estos coeficientes engendran una ampliación $K = R(u_0, u_1, \ldots, u_{n-1})$, la cual es una ampliación finita del campo R de los racionales, por el Corolario 4 del Teorema 10. Cualquier raíz compleja r de la ecuación dada es algebraica sobre el campo K, así que $K(r)$ es una extensión finita de K y, por lo tanto, de R. Cualquier elemento r de esta ampliación es algebraico sobre R, por el Teorema 9. Esto significa que la raíz r es un número algebraico en el campo A, así que A es algebraicamente cerrado.

Tenemos ahora el campo racional R sumergido en el campo algebraicamente cerrado A de los números algebraicos, y el campo R^*

(*) El autor declara preferible, por razones de analogía con la topología, la locución de «algebraicamente completo» que la de «algebraicamente cerrado», adoptada en los libros clásicos y que es también la usual en español, por cuyo motivo la adoptamos en esta traducción. (N. del T.)

de los números reales sumergido en el campo algebraicamente cerrado de los números complejos. Estos resultados son casos particulares de un teorema que asegura, que para cualquier campo F hay una ampliación algebraicamente cerrada A cuyos elementos son todos algebraicos sobre F (cfr. Cap. XV, §1, Apéndice).

La teoría de los números algebraicos ha sido desarrollada muy ampliamente. Se refiere principalmente a campos K de números algebraicos que son ampliaciones finitas del campo R. A los tales se les llama *campos de números algebraicos*. Ahora pasamos a considerar las propiedades aritméticas de estos campos.

EJERCICIOS

1. Ilustrar el Teorema 12, hallando una ecuación con coeficientes racionales para cada uno de los números siguientes:
   a) $\sqrt{2}+\sqrt{-3}$;
   b) $\sqrt{-1}+\sqrt[3]{5}$;
   c) $(\sqrt{7})(\sqrt[3]{2})$;
   d) $\sqrt{7}/(1+\sqrt{2})$;
   e) $u\sqrt{-2}$, con $u^3+7u-14=0$.
2. a) Si u y v son algebraicos de grados m y n, respectivamente (sobre R), demostrar que el grado de $u+v$ no excede a mn.
   b) ¿Cuál es el grado máximo de u/v?
   c) Si t es trascendente y u algebraico, probar que $t+u$ y tu son trascendentes, excepto, en el último caso, cuando $u=0$.
3. Ilustrar el Teorema 13 encontrando una ecuación con coeficientes enteros para una raíz de las siguientes ecuaciones:
   a) $x^2+3x+\sqrt{2}=0$;
   b) $x^2+\sqrt{3}x-\sqrt{-1}=0$;
   c) $x^3-\sqrt{3}x+1+\sqrt[3]{2}=0$;
   d) $x^3+u+2=0$, donde u es una raíz de $u^5+5u^3-10u+5=0$.
4. Presentar los seis primeros polinomios cuadráticos en la alineación de los polinomios cuadráticos del Teorema 11.
5. Demostrar que el conjunto de todos los números algebraicos de determinado grado es numerable, sin utilizar el Teorema 11.
6. Demostrar que cualquier ampliación finita de un campo numerable es numerable.
7. Demostrar que el conjunto A de todos los elementos de un campo F que son algebraicos sobre un subconjunto numerable de F, es numerable.
8. a) Demostrar que existen números reales trascendentes sobre $R(\pi)$.
   b) Utilizando el Ejercicio 7, demostrar que existe una multitud numerable de números reales algebraicamente independientes.

9. Mostrar que la demostración dada del Teorema 11 utiliza implícitamente las siguientes fórmulas de aritmética transfinita: a) Existen $d^{n+1}=d$ polinomios de grado n; b) Existen $d+d+\dots+d+\dots$ (d sumandos)$=d^2$ polinomios de todos los grados.

*10. a) Si u es un número fijo real, mostrar por factorización de x^j-u^j que existe una constante N tal, que $|x^j-u^j|\leqslant N|x-u|$, siempra que $|x-u|<1$.

b) Si $f(x)$ es un polinomio con coeficientes reales, y u un número real, demostrar que existe una constante M que depende de f y de u tal, que $|f(x)-f(u)|\leqslant M|x-u|$, siempre que $|x-u|<1$.

*11. Sea u una raíz real de la ecuación polinómica $f(x)=0$ de grado r, con coeficientes enteros. Si m y n son dos enteros tales, que $|m/n-u|<1/Mn^r$, donde M es la constante del Ejercicio 10, mostrar que $f(m/n)=0$. (*Sugerencia:* Por el Ejercicio 10, $|f(m/n)|<1/n^r$, mientras que $f(m/n)$ es un número racional de denominador n^r.)

*12. Si u es un número real para el cual puede hallarse una sucesión infinita de distintas fracciones racionales m_k/n_k tales, que $|u-(m_k/n_k)|<1/kn_k^k$ para todo k, demostrar que u es trascendente. (*Sugerencia:* Si el grado de u fuese r, por Ejercicio 11 sería $f(m_k/n_k)=0$ para k suficientemente grande.)

*13. Los números que satisfacen las hipótesis del Ejerc. 12 son llamados «números (trascendentes) de Liouville». a) Demostrar que $\sum_{k=1}^{\infty}10^{-k!}=0{,}110001\dots$ es un número de Liouville. b) Presentar otros números de Liouville.

7. Enteros de Gauss

Un ejemplo simple de enteros algebraicos es el que ofrecen los *enteros de Gauss*. Se llama así a los números $\alpha=a+bi$ en $R(i)$ que tienen sus componentes a y b enteros, es decir, en J. La suma, diferencia y producto de dos de estos enteros es también un entero, de modo que los enteros de Gauss forman un dominio de integridad $J[i]$. En este dominio puede atenderse a las cuestiones de divisibilidad y descomposición en factores primos (irreducibles).

Es conveniente introducir la «norma» de un complejo σ cualquiera (entero o no). Si $\sigma=r+si$, la *norma* $N(\sigma)$ es, por definición, el producto de σ por su conjugado $\sigma^*=r-si$,

$$N(\sigma)=\sigma\sigma^*=(r+si)(r-si)=r^2+s^2. \tag{10}$$

La norma no puede ser negativa jamás, y es el cuadrado del valor absoluto de σ (ver Cap. V). Para dos números cualesquiera σ y τ se tiene

$$N(\sigma\tau)=N(\sigma)N(\tau). \tag{11}$$

Esta igualdad indica que la correspondencia $\sigma \to N(\sigma)$ conserva los productos, es decir, que es una representación homomorfa del grupo multiplicativo de los números σ no nulos sobre el grupo multiplicativo de los números racionales. En particular, la norma de un entero gaussiano es un entero (racional).

Recordemos ahora los conceptos generales que implica la divisibilidad (Cap. IV, § 5). Una *unidad* de $J[i]$ es un entero gaussiano $\alpha \neq 0$ cuyo recíproco α^{-1} es también un entero gaussiano. Como $\alpha\alpha^{-1}=1$, resulta $N(\alpha\alpha^{-1})=N(\alpha)N(\alpha^{-1})=1$, así que la norma de una unidad α debe ser $N(\alpha)=1$. La inspección de (10) muestra inmediatamente que las únicas unidades posibles son ± 1 y $\pm i$. Dos enteros se llaman *asociados* en $J[i]$ si cada uno divide al otro. Por lo tanto, los únicos asociados de α en $J[i]$ son $\pm\alpha$ y $\pm i\alpha$.

El número racional primo 5 tiene en $J[i]$ cuatro descomposiciones factoriales posibles:

$$(12) \qquad 5=(1+2i)(1-2i)=(2i-1)(-2i-1)= \\ =(2+i)(2-i)=(i-2)(-i-2).$$

Estas descomposiciones no son esencialmente diferentes; por ejemplo, $2+i=i(1-2i)$ y $2-i=-i(1+2i)$, y en cada uno de los otros casos, los factores correspondientes son asociados. Cada factor en (12) es *primo* (irreducible). Por ejemplo, si $2+i$ tuviese una factorización $2+i=\alpha\beta$, sería $N(2+i)=5=N(\alpha)N(\beta)$, así que $N(\alpha)$ [o $N(\beta)$] sería igual a 1, es decir, que α (o β) sería una unidad. Los factores (12) dan, esencialmente, las únicas maneras de descomponer factorialmente el 5, pues en cualquier descomposición $5=\gamma\delta$, será $N(5)=25=N(\gamma)N(\delta)$, de modo que todo factor que no sea unidad deberá tener norma 5. Pero por tanteos sencillos en (10), se comprueba inmediatamente que los únicos enteros de norma 5 son los utilizados en (12).

Por otra parte, el primo racional 3 es primo en $J[i]$. Para demostrarlo, supongamos $3=\alpha\beta$. Entonces, $N(\alpha)N(\beta)=9$ y $N(\alpha)\mid 9$. Si $N(\alpha)=1$, α es una unidad, mientras que si $N(\alpha)=N(a+bi)=3$, será $a^2+b^2=3$, lo cual es imposible siendo a y b enteros.

El teorema de factorización única puede demostrarse para los enteros de Gauss, desarrollando primero un algoritmo de la división análogo al utilizado con los enteros ordinarios y con los polinomios.

TEOREMA 14. *Dados dos enteros gaussianos α y β, existen otros dos enteros gaussianos γ y ρ cumpliendo las condiciones*

$$(13) \qquad \alpha=\beta\gamma+\rho, \quad N(\rho)<N(\beta).$$

Demostración. Calculemos el cociente exacto $\alpha/\beta=r+si$, y tomemos los dos enteros r' y s' lo más próximos a r y s. Se tiene entonces

$$\alpha/\beta=(r'+s'i)+[(r-r')+(s-s')i]=\gamma+\sigma, \quad \gamma=r'+s'i,$$

donde $|r-r'|\leqslant 1/2$, $|s-s'|\leqslant 1/2$, así que

$$N(\sigma)=(r-r')^2+(s-s')^2\leqslant 1/4+1/4<1.$$

La igualdad (13) puede ahora escribirse $\alpha=\beta\gamma+\beta\sigma$, donde α y $\beta\gamma$ son enteros, y por tanto, también lo es $\beta\sigma$, teniéndose además $N(\beta\sigma)=N(\beta)N(\sigma)<N(\beta)$, c. q. d.

LEMA 1. *Dos enteros gaussianos α_1 y α_2 tienen un máximo común divisor δ, el cual es un entero gaussiano expresable en la forma $\delta=\beta_1\alpha_1+\beta_2\alpha_2$, siendo β_1 y β_2 otros dos enteros gaussianos.*

Demostración. Por divisiones reiteradas se puede construir un algoritmo de Euclides, exactamente como en el caso de los enteros racionales (Cap. I, § 7). Los sucesivos restos ρ de (13) decrecen en norma y, por lo tanto, el algoritmo debe tener un final. El último resto distinto de cero es el m. c. d. buscado.

Una demostración menos directa puede resultar de la consideración del ideal (α_1, α_2) engendrado por α_1 y α_2 en el anillo $J[i]$. Entre los elementos de este ideal elijamos uno, δ, de norma mínima, y escribamos $\alpha_1=\delta\gamma_1+\rho_1$, $\alpha_2=\delta\gamma_2+\rho_2$, como en (13). Los restos ρ_i están en el ideal, y tienen norma menor que δ, luego deben ser cero. Por lo tanto, $\alpha_1=\delta\gamma_1$, $\alpha_2=\delta\gamma_2$, así que δ es un divisor común. Como δ está en el ideal, debe ser de la forma $\delta=\beta_1\alpha_1+\beta_2\alpha_2$ y, por lo tanto, será múltiplo de cualquier divisor común de α_1 y α_2. Por lo tanto, δ es el m. c. d. requerido.

El estudio de la factorización en los enteros gaussianos es completamente análogo al del caso de los enteros racionales (Cap. I, §§ 7-8) y al de los polinomios (Cap. IV, §§ 6-7). Por lo tanto, mencionaremos sólo los puntos más importantes. Un entero de Gauss π se llama primo si no es cero ni unidad y si sus únicos divisores en $J[i]$ son las unidades y sus asociados. Se demuestra :

LEMA 2. *Si π es primo, la relación $\pi \mid \alpha\beta$ implica que $\pi \mid \alpha$ o que $\pi \mid \beta$.*

TEOREMA 15. *Cualquier entero de Gauss α puede expresarse como un producto $\alpha = \pi_1 \ldots \pi_n$ de factores enteros gaussianos primos. Esta representación es esencialmente única, en el sentido de que cualquier otra descomposición de α en factores primos tiene el mismo número de ellos y pueden ordenarse de modo que los factores homólogos sean asociados.*

Para generalizar de modo apropiado estas nociones, investigaremos primero las ecuaciones polinómicas irreducibles satisfechas por los enteros gaussianos. Si $\alpha = a + bi$ es uno de los tales, que no sea un entero racional, será $b \neq 0$, y α debe ser raíz del polinomio cuadrático irreducible siguiente:

$$[x - (a + bi)][x - (a - bi)] = x^2 - 2ax + (a^2 + b^2).$$

Así, α es raíz de una ecuación mónica irreducible con coeficientes enteros. Inversamente, puede demostrarse que si un número $r + si$ en el campo $R(i)$ satisface a una ecuación mónica irreducible con coeficientes enteros, tal número es un entero de Gauss (*). Así resulta:

TEOREMA 16. *Un número del campo* R(i) *es un entero de Gauss si, y sólo si, la ecuación mónica irreducible a que satisface sobre* R *tiene coeficientes enteros.*

EJERCICIOS

1. Descomponer en factores primos los siguientes enteros de Gauss: 5, $3+i$, $6i$, 11, $1-7i$.
2. Hallar el m. c. d. de los siguientes pares de números α_1 y α_2 y expresarlo en la forma $\beta_1\alpha_1 + \beta_2\alpha_2$:
   a) $3+6i$ y $12-3i$; b) $5+3i$ y $13+18i$.
3. Hallar todas las factorizaciones posibles de 13 en enteros gaussianos y mostrar explícitamente que dos factorizaciones difieren sólo en factores asociados.
4. Demostrar que cualquier ideal de $J[i]$ es principal.
5. a) Demostrar el Lema 1 utilizando el algoritmo euclídeo.
   b) Demostrar el Lema 2.

(*) La demostración será desarrollada en un caso más general en la próxima sección (Teorema 18).

6. Demostrar el Teorema 15 a partir del Lema 2.
7. a) Demostrar que un número racional primo p es primo en $J[i]$ si, y sólo si, $x^2+y^2=p$ no tiene soluciones enteras.
   b) Demostrar que todo primo racional de la forma $p=4n+3$ es primo en $J[i]$.
8. a) Demostrar que el anillo cociente $J[x]/(p, x^2+1)$ es isomorfo con $J[i]/(p)$ y con $J_p[x]/(x^2+1)$.
   b) Demostrar que el primero es un dominio de integridad si, y sólo si, p es primo en $J[i]$; mientras que el segundo es un dominio de integridad si, y sólo si, $x^2\equiv -1$ (mód. p) no tiene solución en J.
   c) Admitiendo que el grupo multiplicativo mód. p es cíclico (Cap. XV, Teorema 18), demostrar que si $p=4n+1$, hay en J una solución de $x^2\equiv -1$ (mód. p).
   d) Concluir que $p=4n+1$ no puede ser primo en $J[i]$.

Los ejercicios que siguen se refieren al dominio de todos los números $a+b\sqrt{2}$, donde a y b son enteros.

9. Definida la norma como $N(a+b\sqrt{2})=a^2-2b^2$, mostrar sus propiedades.
10. Establecer un algoritmo de división en $J[\sqrt{2}]$.
11. Demostrar la existencia de m.c.d. en $J[\sqrt{2}]$.
12. Enunciar y demostrar el teorema de descomposición factorial única en $J[\sqrt{2}]$.
13. Hallar en $J[\sqrt{2}]$ la descomposición de los siguientes números: 5, $2+\sqrt{2}$, $1+3\sqrt{2}$.
14. a) Encontrar en $J[\sqrt{2}]$ una unidad distinta de ± 1.
    b) Demostrar que existe en $J[\sqrt{2}]$ un número infinito de unidades distintas. (*Sugerencia:* Utilizar las potencias de una unidad.)

9. Enteros algebraicos

En general, un número algebraico u se llama *entero algebraico* si la ecuación *mónica* irreducible satisfecha por u sobre el campo de los números racionales tiene enteros sus coeficientes; esto es, si

$$(14)\qquad p(u)=a_0+a_1u+\ldots+a_{n-1}u^{n-1}+u^n=0 \quad (a_i \text{ enteros}),$$

donde $p(x)$ es irreducible sobre R. La ecuación irreducible satisfecha por el número racional m/n es precisamente la ecuación lineal $x-m/n=0$. Por lo tanto, un número racional es un entero algebraico si, y sólo si, es un entero ordinario. Tal entero (ordinario) de J será llamado *entero racional* para distinguirlo de los enteros algebraicos. Un número algebraico $u\neq 0$ es llamado *unidad* cuando tanto u como u^{-1} son enteros algebraicos.

Para averiguar si un número algebraico dado es entero, no es necesario acudir a la ecuación irreducible, como demuestra el siguiente resultado:

TEOREMA 17. *Un número es un entero algebraico si, y sólo si, satisface sobre* R *a una ecuación polinómica mónica y con coeficientes enteros.*

Demostración. Supongamos que u es raíz de algún polinomio $f(x)$ mónico con coeficientes enteros. Sobre R, u anula también a un polinomio irreducible $p(x)$, el cual puede tomarse con coeficientes enteros. Suprimiendo cualquier factor común de estos coeficientes, se logrará que el m. c. d. de todos ellos sea 1. Esto es como decir que $p(x)$ es primitivo, en el sentido del Cap. IV, §8, en el dominio $J[x]$ de todos los polinomios con coeficientes enteros. El polinomio dado $f(x)$ es mónico, luego también es primitivo. Por el Teor. 2 sabemos que el polinomio $f(x)$ con raíz u debe ser divisible, en $R[x]$, por el polinomio irreducible $p(x)$ correspondiente a u; y así $f(x)=p(x)q(x)$. Como f y p son primitivos, el Lema 8, §8, Cap. IV, asegura que el cociente $q(x)$ tiene también coeficientes enteros. El coeficiente principal en $f(x)$, 1, es entonces el producto de los coeficientes principales en q y en p; por lo tanto, $\pm p(x)$ es mónico, lo cual significa que u es entero de acuerdo con la definición (14).

Un número puede ser entero algebraico aunque su aspecto sea fraccionario; por ejemplo, $u=(1+\sqrt{5})/2$ tiene forma de fracción, pero satisface a la ecuación

$$[x-(1+\sqrt{5})/2][x-(1-\sqrt{5})/2]=x^2-x-1=0,$$

la cual es mónica y de coeficientes enteros. Esto nos incita a una investigación sistemática de los números de un campo cuadrático que sean enteros algebraicos. Cualquier campo K de grado 2 sobre el campo R de los racionales puede ser expresado como una ampliación algebraica simple $K=R(\sqrt{d})$. Sin pérdida de generalidad se puede suponer que d es un entero racional y que no admite como factor ningún cuadrado de otro entero. En este caso, resulta:

TEOREMA 18. *Si* $\mathrm{d}\neq 1$ *es un entero sin factores cuadrados, se verifica: Si* $\mathrm{d}\equiv 2$ o $\mathrm{d}\equiv 3$ *(mód. 4), los enteros algebraicos en* $\mathrm{R}(\sqrt{\mathrm{d}})$ *son los números* $\mathrm{a}+\mathrm{b}\sqrt{\mathrm{d}}$, *con los coeficientes* a *y* b *enteros racio-*

nales. Si es $d \equiv 1$ *(mód. 4), los enteros de* $R(\sqrt{d})$ *son los números* $a+b(1+\sqrt{d})/2$*, con* a *y* b *enteros racionales.*

Demostración. Como preliminar, observemos que $a \equiv 1$ (mód. 2) significa que $a=1+2r$, es decir, que $a^2=1+4r+4r^2 \equiv 1$ (mód. 4). En otras palabras:

(15) $a \equiv 1$ (mód. 2) implica $a^2 \equiv 1$ (mód. 4),

(16) $a \equiv 0$ (mód. 2) implica $a^2 \equiv 0$ (mód. 4),

así que un cuadrado es siempre congruente con 0 o con 1, módulo 4.

Cualquier número u en $R(\sqrt{d})$ puede expresarse como $u = (a+b\sqrt{d})/c$, donde los enteros a, b y c no tienen ningún factor común. Supondremos $b \neq 0$, para excluir el caso trivial de los números racionales. La ecuación cuadrática mónica irreducible para u es la siguiente:

(17) $$[x-(a+b\sqrt{d})/c][x-(a-b\sqrt{d})/c] = x^2-(2a/c)x+(a^2-db^2)/c^2=0.$$

Si u es un entero algebraico, los coeficientes $2a/c$ y $(a^2-db^2)/c^2$ deberán ser enteros. Por lo tanto, $4a^2/c^2$, $(4a^2-4db^2)/c^2$ y $4db^2/c^2$ deberán también ser enteros, así que $c \mid 2a$ y $c^2 \mid 4db^2$. Como d no tiene factores cuadrados, cualquier primo $p \neq 2$ contenido en c debe dividir a a y a b^2, contra el supuesto de que a, b y c no tienen factores comunes (excepto ± 1). Por la misma razón, $4 \mid c$ es imposible, así que las solas posibilidades son $c=1$ y $c=2$.

Consideremos ahora el caso $d \equiv 2$ o $d \equiv 3$ (mód. 4), con $c=2$. En este caso, el último coeficiente $(a^2-db^2)/4$ de (17) debe ser entero, y por lo tanto, $a^2 \equiv db^2$ (mód. 4). Si $b \equiv 1$ (mód. 2), será $b^2 \equiv 1$ (módulo 4) y $a^2 \equiv db^2 \equiv 2$ ó 3 (mód. 4), en contradicción a las reglas (15) y (16). Si $b \equiv 0$ (mód. 2), será $a^2 \equiv 0$ (mód. 4) y $a \equiv 0$ (mód. 2), así que a, b y c tienen como factor común el 2. En cualquier caso concluímos que no puede ser 2; y por ser $c=1$, todos los enteros de $R(\sqrt{d})$ serán de la forma $a+b\sqrt{d}$. Inversamente, la ecuación mónica (17) para los números de este tipo tiene coeficientes enteros.

Para el otro caso, $d \equiv 1$ (mód. 4), vale una demostración análoga, observándose ahora la posibilidad de $a \equiv b \equiv 1$ (mód. 2).

COROLARIO. *En cualquier campo de grado 2 sobre R, el conjunto de todos los enteros algebraicos es un dominio de integridad.*

Demostración. Las sumas, diferencias y productos de los enteros en la forma expresada en el Teorema 18, son también números de la misma forma, y por ende enteros.

Inmediatamente vamos a generalizar este corolario a cualquier campo de números algebraicos.

EJERCICIOS

1. Demostrar que cualquier raíz de la unidad es un entero algebraico.
2. a) Hallar todos los enteros y todas las unidades en $R(\omega)$, donde ω es una raíz cúbica compleja de la unidad.
   b) Probar que cualquier unidad en $R(\omega)$ es una raíz de la unidad.
3. Completar la demostración del segundo caso del Teor. 18 [$d \equiv 1$ (mód. 4)].
4. a) Demostrar que cualquier número algebraico puede ser escrito como un cociente u/b, donde u es un entero algebraico y b un entero racional (es decir, un entero de J).
   b) Demostrar que cualquier campo K de números algebraicos es el campo de cocientes del dominio de todos los enteros algebraicos en K.
*5. Hallar todos los enteros en $R(\sqrt{2}, i)$.

9. Sumas y productos de enteros

Esta sección se dedica a la demostración del siguiente resultado:

TEOREMA 19. *El conjunto de todos los enteros algebraicos es un dominio de integridad.*

Una consecuencia inmediata es la siguiente particularización:

COROLARIO. *En cualquier campo* K *de números algebraicos, los enteros algebraicos forman un dominio de integridad.*

Una instructiva demostración del Teorema 19 se deduce del análisis de los grupos aditivos engendrados por enteros algebraicos. Si $v_1, \ldots, v_n$ son números algebraicos cualesquiera, denotemos con $G = [v_1, \ldots, v_n]$ el subgrupo (*) engendrado por estos números en el grupo aditivo de los números complejos. Este grupo G consiste, sencillamente, en todos los números representables en la forma

(18) $$u = a_1v_1 + a_2v_2 + \ldots + a_nv_n \quad (a_i \text{ entero racional}).$$

Observemos que en el sentido de grupo aditivo, el múltiplo natural $av = a \times v$ es, simplemente, una potencia de v.

(*) Un grupo aditivo y abeliano, como lo es G, se llama a veces *módulo* o *J-módulo.*

LEMA 1. *Cualquier subgrupo* S *del grupo* $G=[v_1, \ldots, v_n]$ *puede también ser engendrado por* n *o menos números.*

Demostración. Para cada índice k, sea G_k el subgrupo $[v_k, \ldots, v_n]$ engendrado por los últimos $n-k+1$ generadores de G; así que G_k consiste en todas las sumas de la forma $a_kv_k+\ldots+a_nv_n$. Entre todos los elementos de G_k que pertenezcan al subgrupo dado S, escojamos un elemento

$$(19) \qquad w_k=c_kv_k+c_{k+1}v_{k+1}+\ldots+c_nv_n,$$

en el cual el primer coeficiente c_k tenga el valor positivo más pequeño posible. (Si en cada elemento el coeficiente de v_k es cero, pondremos $w_k=0$.) Si $w=b_kv_k+\ldots$ es otro elemento de S en G_k, su primer coeficiente b_k podrá escribirse $b_k=q_kc_k+r_k$, con un resto no negativo $r_k<c_k$. La diferencia $w-q_kw_k=r_kv_k+\ldots$ pertenece entonces al subgrupo G_k y a S, y tiene un primer coeficiente no negativo $r_k<c_k$. Por lo tanto, $r_k=0$, y cualquier elemento w de S en G_k dará un elemento $w'=w-q_kw_k$ en G_{k+1}.

Los n elementos escogidos $w_1, \ldots, w_n$ engendran la totalidad del grupo S. Puesto que, dado un elemento w en el grupo, se puede hallar un q_1 tal, que $w-q_1w_1$ dependa sólo de $v_2, \ldots, v_n$, y luego un q_2 tal, que $w-q_1w_1-q_2w_2$ dependa sólo de $v_3, \ldots, v_n$, y así sucesivamente; y al final, $w=\Sigma q_iw_i$. Esto demuestra el lema.

El método utilizado en el Lema 1 es empleado asimismo en muchas cuestiones referentes a los grupos abelianos con un número finito de generadores.

LEMA 2. *Un número* u *es un entero algebraico si, y sólo si, el grupo aditivo engendrado por todas las potencias* 1, u, u^2, u^3, ... *de* u, *puede ser engendrado también por un número finito de elementos.*

Demostración. Si u es un entero, satisface a una ecuación mónica (14) de grado n con coeficientes enteros. Esta ecuación expresa u^n como un elemento del grupo $G=[1, u, \ldots, u^{n-1}]$ engendrado por las n primeras potencias de u. Por iteración, la misma ecuación puede utilizarse para expresar cualquier potencia superior de u como un elemento de este grupo. Por lo tanto, u satisface al criterio del Lema 2.

Inversamente, supongamos que el grupo G engendrado por 1, u, u^2, ... puede ser engendrado por los n números v_1, ..., v_n pertenecientes a G. El producto de u por un elemento $\Sigma a_j u^j$ de G es también un elemento $\Sigma a_j u^{j+1}$ de G, así que los productos uv_i deben pertenecer a G y ser expresables mediante los generadores como $uv_i = \sum_j a_{ij} v_j$, donde los a_{ij} son enteros. Esta expresión proporciona n ecuaciones homogéneas en las v, de la forma

$$\begin{aligned} (a_{11} - u)v_1 + a_{12}v_2 + \ldots + a_{1n}v_n &= 0, \\ a_{21}v_1 + (a_{22} - u)v_2 + \ldots + a_{2n}v_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}v_1 + a_{n2}v_2 + \ldots + (a_{nn} - u)v_n &= 0. \end{aligned}$$

Este sistema de ecuaciones tiene un conjunto de soluciones no todas nulas, v_1, v_2, ..., v_n, luego la matriz de los coeficientes debe ser singular (Cap. X, Teor. 8, Corolario). La matriz de los coeficientes puede escribirse como $A - uI$, donde $A = \| a_{ij} \|$; como es singular, su determinante es cero, así

(20) $$|A - uI| = (-1)^n u^n + b_{n-1} u^{n-1} + \ldots + b_n = 0,$$

donde los coeficientes b_i son ciertos polinomios en los enteros a_{ij}, luego son también enteros. Esta ecuación (20) significa (*) que u es un entero algebraico, como afirmaba el lema.

La conclusión del Lema 2 puede ser formulada así:

Corolario. *Si todas las potencias positivas de un número algebraico* u *pertenecen al grupo aditivo engendrado por un conjunto finito de números* y_1, ..., y_n, *entonces* u *es un entero algebraico.*

Demostración. El grupo S engendrado por 1, u, u^2, ... es un subgrupo del grupo engendrado por 1, y_1, ..., y_n. Por lo tanto, según el Lema 1, este subgrupo S puede ser engendrado por un número finito de sus elementos y, por lo tanto, según el Lema 2, el número u es un entero algebraico.

Volvamos ahora a la demostración del Teorema 19. Si u y v son enteros algebraicos, vamos a probar que $u+v$ y uv son enteros. La hipótesis significa que todas las potencias u^k y v^k pueden ser expre-

(*) Obsérvese que (20) es, simplemente, el polinomio característico de A, en el sentido de Cap. X.

sadas mediante un número finito de potencias $1, u, \ldots, u^{n-1}$ y $1, v, \ldots, v^{r-1}$. Por lo tanto, cualquier potencia $(uv)^k = u^k v^k$ y $(u+v)^k$ pertenece al grupo aditivo engendrado por los productos $1, u, uv, uv^2, \ldots, u^{n-1}v^{r-1}$. Y por el corolario anterior, esto demuestra que uv y $u+v$ son enteros algebraicos, como afirma el teorema.

EJERCICIOS

1. Plantear la oportuna ecuación mónica irreducible con coeficientes enteros que muestre que cada uno de los siguientes números es un entero algebraico:

   a) $\sqrt{2}+\sqrt{3}$; b) $i+\omega$; c) $\sqrt{7}+(1+\sqrt{5})/2$.

2. a) Si los números $v_1, \ldots, v_n$ son linealmente independientes sobre R, demostrar que cualquier subgrupo S de índice finito en $G=[v_1, \ldots, v_n]$ puede también ser engendrado por n números linealmente independientes $w_1, \ldots, w_n$.

   b) Mostrar que este subgrupo S es isomorfo con el grupo completo G.

3. Si los números $v_1, \ldots, v_n$ son linealmente independientes sobre R, mostrar de qué modo la base establecida en el lema 1 para un subgrupo S de $G=[v_1, \ldots, v_n]$ puede utilizarse para calcular el índice de S en G. (*Sugerencia:* Hallar primero un elemento representativo para cada clase de restos determinada por S.)

*4. Demostrar que un grupo $G=[v_1, \ldots, v_n]$ no tiene una cadena infinita de subgrupos distintos, es decir, que, dada una sucesión infinita de subgrupos $S_1 \leqslant S_2 \leqslant S_3 \leqslant \ldots \leqslant G$, hay un índice m para el cual $S_m = S_{m+1} = S_{m+2} = \ldots$ (*Sugerencia:* Aplicar el Lema 1 a la reunión de los grupos S_k.)

5. a) Demostrar que cualquier módulo contenido en el dominio J de los enteros ordinarios, es un ideal de J.

   b) Mostrar un módulo contenido en el dominio $J[i]$ de los enteros de Gauss, que no sea un ideal de $J[i]$.

*6. Si un número algebraico u es raíz de un polinomio mónico irreducible, cuyos coeficientes son enteros *algebraicos*, demostrar que u es también entero algebraico.

*10. Factorización en los campos cuadráticos

Al estudiar la divisibilidad de los enteros en cualquier campo cuadrático, es siempre conveniente el emplear las normas. La fórmula para la norma depende del campo, pero la idea es la misma en todos los casos, incluso para campos algebraicos de orden superior. La norma se define, esencialmente, mediante los automorfismos del campo. El campo cuadrático $R(\sqrt{d})$ tiene, por el Teorema 7, un isomorfismo $u = a + b\sqrt{d} \leftrightarrow \bar{u} = a - b\sqrt{d}$, en el que a cada número u le corresponde su conjugado $\bar{u}$.

DEFINICIÓN. *La norma* N(u) *del número* $u=a+b\sqrt{d}$ *de* $R(\sqrt{d})$ *es el producto* $u\bar{u}$ *de* u *por su conjugado* $\bar{u}$,

(21) $$N(u)=u\bar{u}=(a+b\sqrt{d})\,(a-b\sqrt{d}).$$

Como la correspondencia $u\leftrightarrow\bar{u}$ es un isomorfismo, $\overline{uv}=\bar{u}\,\bar{v}$, y por lo tanto,

(22) $$N(uv)=N(u)N(v).$$

De este modo, la norma traslada la factorización $w=uv$ de cualquier entero del campo, a la factorización $N(w)=N(u)N(v)$ de los enteros racionales $N(w)$. (La norma de un entero algebraico es un entero racional; ver Ejerc. 1.)

Para probar que la descomposición factorial en números primos no es siempre única, consideremos el campo $R(\sqrt{-5})$. De acuerdo con el Teorema 18, cualquier entero de este campo tiene la forma $a+b\sqrt{-5}$, con a y b enteros racionales; por lo tanto, todos los enteros forman un dominio $J[\sqrt{-5}]$. Si u es una unidad de este dominio, entonces $uu^{-1}=1$, donde u^{-1} es también entero. Además, $N(uu^{-1})=N(u)N(u^{-1})=1$, así que el entero $N(u)$ debe ser ± 1. Pero

(23) $$N(u)=N(a+b\sqrt{-5})=a^2+5b^2.$$

Este valor es ± 1 si, y sólo si, $b=0$ y $a=\pm 1$, así que las únicas unidades de $J[\sqrt{-5}]$ son ± 1.

¿Cómo es la descomposición del entero 21 en $J[\sqrt{-5}]$? Si $21=uv$, para dos enteros u y v, entonces $N(uv)=N(u)N(v)=N(21)=21^2=3^2\cdot 7^2$. Así, un factor propio de 21, sea u, tendrá por norma un factor propio de $3^2\cdot 7^2$, luego $N(u)=3$, 7, 9, 21, 49, 63 ó 147. Como la norma está dada por (23), se encuentran con pocos ensayos todos los enteros posibles con tales normas. Resulta que no hay enteros de norma 3 ó 7 (y por tanto, tampoco factores de normas 63 ó 147); las restantes soluciones de (23) dan las factorizaciones siguientes:

(24) $$21=3\cdot 7=(1+2\sqrt{-5})\,(1-2\sqrt{-5})=(4+\sqrt{-5})\,(4-\sqrt{-5}).$$

Aquí figuran todas las factorizaciones de 21 (excepto las alteraciones por factores unidad ± 1), y cada uno de los factores que aparece es primo en $J[\sqrt{-5}]$. Por ejemplo, 3 es primo, ya que cualquiera de sus eventuales factores primos tendría norma 3 y no po-

dría ser entero. El ejemplo de las diferentes descomposiciones (24) demuestra que en el dominio de los enteros $J[\sqrt{-5}]$ no vale el teorema de factorización única.

La dificultad esencial en este dominio es la imposibilidad de obtener un máximo común divisor, ni por el algoritmo de Euclides ni por factores primos como 3 y $1+2\sqrt{-5}$. Recordaremos que para los enteros de Gauss, el m. c. d. de α_1 y α_2 venía definido como un elemento que por sí solo engendraba el ideal (α_1, α_2) (ver la demostración del Lema 1 en § 7), y que el mismo ideal engendraba el m. c. d. de dos enteros racionales (ver Capítulo XIII, § 4). Este ideal puede también ser formado, en el caso actual, como ideal $(3, 1+2\sqrt{-5})$ engendrado por 3 y $1+2\sqrt{-5}$. Consideremos asimismo los ideales

$$P_1=(3,\ 1+2\sqrt{-5}),\qquad P_2=(3,\ 1-2\sqrt{-5}),$$
$$Q_1=(7,\ 1+2\sqrt{-5}),\qquad Q_2=(7,\ 1-2\sqrt{-5}). \tag{25}$$

Dos de estos ideales pueden multiplicarse por simple multiplicación de sus elementos base, por la regla (15) del Cap. XIII; por ejemplo:

$$P_1P_2=(9,\ 3-6\sqrt{-5},\ 3+6\sqrt{-5},\ 1+20).$$

El ideal producto contiene al 9 y al 21, luego contiene a su m. c. d. 3. Los cuatro elementos base de P_1P_2 son múltiplos de 3, así P_1P_2 es simplemente el ideal principal (3) consistente en todos los múltiplos de 3 en $J[\sqrt{-5}]$. De modo semejante se prueba

$$P_1P_2=(3),\quad P_1Q_1=(1+2\sqrt{-5}),\quad P_1Q_2=(4-\sqrt{-5}),$$
$$Q_1Q_2=(7),\quad P_2Q_2=(1-2\sqrt{-5}),\quad P_2Q_1=(4+\sqrt{-5}). \tag{26}$$

Las tres factorizaciones esencialmente distintas que se ven en (24) pueden ahora ser consideradas como factorizaciones del ideal principal engendrado por 21. Sustituyendo los productos (26) en (24) resulta que todas las factorizaciones conducen a una, y a la misma, descomposición factorial de ideales $(21)=P_1P_2Q_1Q_2$. Esto indica que con los *ideales* se reconstruye la unicidad de la descomposición factorial.

Puede objetarse que en la formación de nuestros ideales no hemos utilizado ningún factor $4\pm\sqrt{-5}$ del tercer par (24). Pero es

que cualquier m.c.d. formado a partir de estos factores conduce a los mismos cuatro ideales P_1, P_2, Q_1, Q_2 del caso anterior. Por ejemplo, el ideal $(3, 4-\sqrt{-5})$ contiene a

$$4-\sqrt{-5}-3(1-\sqrt{-5})=1+2\sqrt{-5},$$

e inversamente, el ideal $(3, 1+2\sqrt{-5})$ contiene al elemento $1+2\sqrt{-5}+3(1-\sqrt{-5})=4-\sqrt{-5}$. Por lo tanto,

(27) $$(3, 4-\sqrt{-5})=(3, 1+2\sqrt{-5})=P_1.$$

Los ideales P_1, P_2, Q_1, Q_2 de la factorización que estamos considerando, son ideales primos en el dominio $J[\sqrt{-5}]$. Consideremos el caso de P_1, el cual, por definición, contiene al 3. Demostremos, primero, que los únicos enteros racionales de P_1 son los múltiplos de 3. Pues otros enteros racionales sólo podrían estar en P_1 si lo estuviesen el 1 o el 2; en el último caso, 2 en P_1 implicaría que $1=3-2$ estuviese en P_1. Esto es imposible, pues entonces sería $P_1=(1)$, en cuyo caso P_1 incluiría a todo el anillo $J[\sqrt{-5}]$; mas como el isomorfismo $\sqrt{-5}\to-\sqrt{-5}$ transporta a P_1 sobre P_2, se sigue, por simetría, que también P_2 sería igual todo el anillo. El producto $P_1P_2=(1)(1)=(1)$ no sería entonces el ideal (3), contra lo que fué calculado antes. Es conveniente escribir $u\equiv v$ (mód. P_1) para significar que $u-v$ es un número de P_1. Por (27), P_1 contiene a $4-\sqrt{-5}$, así que $\sqrt{-5}\equiv4\equiv1$ (mód. P_1) y $a+b\sqrt{-5}\equiv a+b$ (mód. P_1). Este resultado significa que cualquier entero algebraico $u=a+b\sqrt{-5}$ es congruente, módulo P_1, a un entero racional $c=a+b$.

Probaremos ahora que P_1 es primo. Por definición, esto significa que

(28) $uv\equiv0$ (mód. P_1) implica $u\equiv0$ (mód. P_1) o $v\equiv0$ (mód. P_1).

Tomemos dos enteros racionales c y d tales, que $u\equiv c$ y $v\equiv d$ (mód. P_1). Entonces $uv\equiv0$ da $cd\equiv0$ (mód. P_1), lo cual implica que $cd\equiv0$ (mód. 3), ya que los únicos enteros racionales de P_1 son los multiplos de 3. Pero 3 es primo, así que $cd\equiv0$ (mód. 3) exige que $c\equiv0$ o $d\equiv0$ (mód. 3), y por lo tanto, $u\equiv0$ o $v\equiv0$ (mód. P_1), como dice (28). Se prueba de modo semejante que P_2, Q_1 y Q_2 son ideales primos.

La descomposición factorial única deducida en el dominio $R(\sqrt{-5})$ es una indicación de cómo la noción de ideal puede ser

empleada sistemáticamente para restablecer el teorema de la descomposición factorial única, en aquellos dominios de números algebraicos donde, con la descomposición factorial ordinaria, la unicidad cae en defecto. Para un desarrollo posterior, se debe establecer el «teorema fundamental de la teoría de ideales» : *En el dominio* D *de todos los enteros algebraicos que pertenecen a un campo* K *de números algebraicos, cualquier ideal puede ser expresado de manera única (salvo el orden) como un producto de ideales primos.*, *En particular, cualquier entero* u *del dominio determina un ideal principal, el cual admite dicha factorización única.*

EJERCICIOS

1. a) Demostrar qué en un campo cuadrático, la norma de un entero es un [illegible]
   b) Si u [illegible] $\sqrt{d}$ no es racional, demostrar que $N(u)$ es el término constante de la ecuación mónica irreducible satisfecha por u.
2. Hallar todas las unidades en $R[\sqrt{-7}]$.
3. Demostrar que, siendo d positivo, el número de unidades en un campo cuadrático $R(\sqrt{-d})$ es finito, y demostrar que cualquier unidad es una raíz de la unidad.
4. Demostrar que todas las raíces de la unidad que pertenecen a cualquier campo de números algebraicos, forman un grupo cíclico.
5. Demostrar con detalle que las únicas factorizaciones de 21 son las dadas en (24).
6. Calcular los productos (26). [*Sugerencia:* Utilizar la simetría; por ejemplo, utilizar el automorfismo de $R(\sqrt{-5})$.]
7. Demostrar que el ideal P_1 del texto no es principal (esto significa que los números 3 y $1+2\sqrt{-5}$ no admiten un número determinado como m. c. d.).
8. Demostrar que cada uno de los ideales $(7, 4-\sqrt{-5})$ y $(1+2\sqrt{-5}, 4+\sqrt{-5})$ es igual a uno de los ideales P_1, P_2, Q_1, Q_2.
9. Demostrar que el ideal Q_1 es primo; de aquí, demostrar por simetría que Q_2 es primo.
10. Demostrar que el anillo cociente $J[\sqrt{-5}]/P_1$ es isomórfico con J_3. (*Sugerencia:* ¿Qué son las congruencias con relación al anillo cociente?)
11. Por el método del Teorema 14, manifestar un algoritmo de división para $J(\sqrt{-2})$.
12. Manifestar un algoritmo de la división para $J(\omega)$, donde $\omega=(-1+\sqrt{-3})/2$. (*Sugerencia:* Los múltiplos enteros de cualquier β dividen al plano complejo en triángulos equiláteros.)
*13. Sea D un dominio de integridad, en el cual una norma $N(\alpha)$ se define así: 1) $N(\alpha)$ es entero positivo si $\alpha\neq 0$; 2) $N(\alpha\beta)=N(\alpha)N(\beta)$; 3) dados α y $\beta\neq 0$, existen γ y ζ tales, que $\alpha=\beta\gamma+\zeta$, $N(\zeta)<N(\beta)$.
   a) Probar que D es un dominio con factorización única.
   b) Probar que cualquier ideal en D es principal.

CAPÍTULO XV

Teoría de Galois

1. Campo raíz de una ecuación

Los complejos conjugados i y $-i$ son las raíces de la ecuación $x^2+1=0$; más generalmente, cualquier par de complejos conjugados $a+bi$ y $a-bi$ puede darse algebraicamente como el par de raíces de la ecuación

$$[x-(a+bi)][x-(a-bi)]=x^2-2ax+(a^2+b^2)=0.$$

Si $b\neq 0$, esta ecuación es irreducible en el campo de los números reales. Análogamente, $\sqrt{2}$ y $-\sqrt{2}$, aunque no complejos conjugados, son llamados *conjugados* sobre el campo racional, ya que son las dos raíces de la ecuación irreducible $x^2-2=0$. Más generalmente, dos elementos u y v de un campo K son conjugados sobre un subcampo F si, y sólo si, u y v son raíces de un mismo polinomio $p(x)$ irreducible sobre F. Por ejemplo, la raíz cúbica real de 5 satisface a la ecuación $x^3=5$, irreducible en el campo racional, y por lo tanto, tiene como conjugados sobre R a los números $\omega\sqrt[3]{5}$ y $\omega^2\sqrt[3]{5}$, donde ω es una raíz cúbica compleja de la unidad.

El estudio de las raíces conjugadas de una ecuación irreducible $p(x)=0$ exige introducir el campo raíz (*), o sea, la ampliación múltiple engendrada por *todas* las raíces sobre el campo de los coeficientes.

(*) *Root field* en el original. Como algunos autores emplean esta denominación para un concepto distinto, quizás hubiese sido conveniente emplear ahora la denominación *campo de descomposición*, que es bastante usual. (N. del T.)

DEFINICIÓN. *Se llama campo raíz de un polinomio* f(x) *de grado* n *con coeficientes en* F, *a una ampliación* N *de* F *tal, que: 1)* f(x) *puede ser descompuesto en factores lineales en* N, $f(x)=c(x-u_1)\ldots(x-u_n)$; *2)* N *es engendrado sobre* F *por las raíces de* f(x), *como* $N=F(u_1, \ldots, u_n)$.

Así pues, este campo raíz será siempre una ampliación finita de F, de grado $n!$ como máximo (cfr. Corol. 4 del Teor. 10 del Capítulo XIV). Por ejemplo, sobre R el polinomio x^3-5 tiene como campo raíz $R(\sqrt[3]{5}, \omega\sqrt[3]{5}, \omega^2\sqrt[3]{5})$, el cual puede también ser engendrado como $R(\sqrt[3]{5}, \omega)$ por los dos números algebraicos $\sqrt[3]{5}$ y ω. El primero satisface a una ecuación de grado tres, y engendra el campo $R(\sqrt[3]{5})$. Sobre este campo, ω satisface a la ecuación ciclotómica $x^2+x+1=0$, que da las raíces cúbicas de la unidad. Esta ecuación es irreducible, por no tener ninguna raíz en $R(\sqrt[3]{5})$; tampoco tiene raíces en la totalidad del campo de los números reales, pues sus dos raíces son complejas. Luego ω tiene grado 2 sobre $R(\sqrt[3]{5})$ y el campo raíz $R(\sqrt[3]{5}, \omega)$ tiene grado 6 sobre R.

Un teorema general de existencia del campo raíz, puede obtenerse a partir del conocido teorema de existencia de las ampliaciones algebraicas simples, como sigue:

TEOREMA 1. *Cualquier polinomio sobre cualquier campo tiene un campo raíz.*

Para un polinomio de primer grado, el campo raíz es precisamente el campo base F; por lo tanto, podemos proceder por inducción, relativa al grado n de $f(x)$. Supongamos el teorema cierto para cualquier campo F y para cualquier polinomio de grado $n-1$, y sea $f(x)$ un polinomio de grado n, siendo $p(x)$ un factor suyo irreducible sobre F. Por el Teorema 6 del Capítulo XIV, existe una ampliación simple $K=F(u)$ engendrada por una raíz u de $p(x)$. Sobre K, $f(x)$ tiene la raíz u, y por consiguiente, un factor $x-u$, así que $f(x)=(x-u)g(x)$. El cociente $g(x)$ es un polinomio de grado $n-1$ y la hipótesis para la inducción proporciona un campo raíz N, engendrado sobre K por las $n-1$ raíces de $g(x)$. Este campo N es un campo raíz para $f(x)$.

Será probado luego (Teor. 6) que todos los campos raíces de un polinomio dado $f(x)$ sobre una determinada base F, son isomorfos; por esto es legítimo decir *el* campo raíz de $f(x)$ sobre F.

Apéndice. Podemos emplear el Teorema 1 en la demostración de la existencia de un campo algebraicamente cerrado, de característica prima cualquiera p, como sigue. La construcción comenzará con el campo J_p de los enteros módulo p. El número de polinomios de grado n sobre J_p es finito, igual a $(p-1)p^n$; por lo tanto, es posible ordenar todos los polinomios sobre J_p en una sucesión infinita $f_1(x)$, $f_2(x)$, $f_3(x)$, ...; en esta sucesión, cualquier polinomio lineal aparecerá antes que cualquier cuadrático, los cuadráticos aparecerán antes que los cúbicos, etc. Definamos ahora el campo F_0 como J_p, y, por recurrencia, definiremos el F_{n+1} como el campo raíz de $f_{n+1}(x)$ sobre F_n. Finalmente, definiremos F^* como el conjunto de todos los elementos que pertenecen a uno o a varios de los campos F_n. Este conjunto F^* es un campo, bajo ciertas operaciones de adición y multiplicación naturalmente definidas; por ejemplo, si a y b son dos elementos de F^*, deberán aparecer ambos en un cierto campo F_n y en todos los siguientes; $a+b$ será el valor común de la suma en este F_n y en sus siguientes. El campo F^* tiene característica p. Para demostrar que F^* es algebraicamente cerrado, sea $g(x)$ cualquier polinomio sobre F^*; todos los coeficientes de $g(x)$ pertenecen a algún F_n, luego son algebraicos sobre J_p. Empleando el Teorema 10, Cap. XIV, se puede hallar un polinomio $h(x)$ múltiplo no nulo de $g(x)$ con coeficientes en J_p (ver el próximo Ejerc. 5). Pero $h(x)$ puede ciertamente ser descompuesto en factores lineales en *su* campo raíz F_m sobre un conveniente F_{m-1}, luego también lo puede ser su divisor $g(x)$. Por lo tanto, $g(x)$ puede ser descompuesto en factores lineales sobre el campo más amplio F^*, el cual es, por lo tanto, un campo algebraicamente cerrado de característica p. Además, cada elemento de F^* es algebraico sobre J_p.

Empleando, en vez de sucesiones, conjuntos bien ordenados en general, y la llamada inducción transfinita, los razonamientos anteriores pueden ser modificados (*) para aplicarlos a cualquier campo F. Esta modificación establece la siguiente generalización parcial del teorema fundamental del Álgebra. *Cualquier campo* F *tiene una ampliación algebraicamente cerrada.*

EJERCICIOS

1. Hallar el grado sobre R del campo raíz de los siguientes polinomios: a) $x^3-x^2-x-2=0$; b) $x^3-2=0$; c) $x^4-7=0$; d) $(x^2-2)(x^2-5)=0$.

(*) Una demostración detallada se encuentra en B. L. van der Waerden, *Moderne Algebra*, I, Berlín, 1930 (en la primera edición, pero no en la segunda).

2. Demostrar: el campo raíz de un polinomio de grado n es, a lo más, $n!$
3. Si ζ es una raíz primitiva n-ésima de la unidad, probar que $R(\zeta)$ es el campo raíz de $x^n-1=0$.
4. En el Apéndice, demostrar con detalle que F^* es un campo.
*5. Sea $g(x)=a_0+a_1x+\dots+a_nx^n$, con coeficientes algebraicos sobre J_p; demostrar que $g(x)$ es divisor de algún $h(x)$ no nulo, con coeficientes de J_p. [*Sugerencia:* Formar un campo raíz para $g(x)$ sobre $J_p(a_0, \dots, a_n)$; descomponer $g(x)$ en factores lineales $(x-u_i)$ en este campo raíz; la u_i será algebraica sobre J_p, definida por el polinomio irreducible $h_i(x)$; poner $h(x)=\prod h_i(x)$.]
*6. Probar (como en el Apéndice) que existe una ampliación algebraicamente cerrada de $J_p(x)$. (*Sugerencia:* Probar el teorema para cualquier campo numerable.)
*7. Demostrar, sin emplear los resultados del Cap. V, que existe una ampliación algebraicamente cerrada del campo racional.

2. El grupo de Galois

Los grupos pueden ser empleados para la valoración de la simetría, no sólo de las figuras geométricas, sino también de los sistemas algebraicos. Por ejemplo, el campo C de los números complejos tiene, con relación al de los reales, dos «simetrías» : una es la identidad y otra es el isomorfismo $a+bi \leftrightarrow a-bi$, el cual representa a cada número sobre su complejo conjugado. Tales isomorfismos, de un campo consigo mismo, son llamados automorfismos. En general, un *automorfismo* T de un campo K es una correspondencia biunívoca $a \leftrightarrow aT$ del conjunto K consigo mismo, de tal modo que las sumas y productos sean conservados, significando esto que

$$(1) \qquad (a+b)T=aT+bT, \qquad (ab)T=(aT)(bT).$$

para todos los a y b en K.

El producto ST de dos automorfismos S y T es también un automorfismo, y la transformación inversa de un automorfismo es también un automorfismo. Por lo tanto,

TEOREMA 2. *Todos los automorfismos de un campo* K *forman un grupo.*

Sea K una ampliación de F y consideremos aquellos automorfismos T tales, que $aT=a$ para todos los a en F ; esto es, los automorfismos T que dejan invariantes los elementos de F. En el grupo de todos los automorfismos de K, estos automorfismos T constituyen

un subgrupo, llamado *grupo de los automorfismos de* K *sobre* F. Así, el grupo de automorfismos de C sobre R^* consiste en los dos automorfismos identidad y conjugación, es decir, $a+bi \leftrightarrow a+bi$ y $a+bi \leftrightarrow a-bi$.

Definición. *El grupo de automorfismos de un campo* K *sobre un subcampo* F *es el grupo de aquellos automorfismos de* K *que dejan invariante cada elemento de* F.

El caso particular más importante lo ofrece el grupo de automorfismos de un campo de números algebraicos sobre el campo R de los racionales, pero antes de considerar ejemplos específicos, determinemos las imágenes posibles de un número algebraico, en un automorfismo dado.

Teorema 8. *Cualquier automorfismo* T *de una ampliación finita* K *sobre* F *hace corresponder a cada elemento* u *de* K *un conjugado de* u *sobre* F, *al que designaremos por* uT.

Este teorema significa que u y su imagen uT deben satisfacer a la misma ecuación irreducible sobre F. En efecto, sea u un elemento dado, algebraico sobre F, raíz del polinomio mónico irreducible $p(x)=x^n+b_{n-1}x^{n-1}+\ldots+b_0$, con coeficientes de F. El automorfismo T conserva todas las relaciones racionales, por (1), y conserva invariante cada b_i; por lo tanto, $p(u)=0$ dará

$$(u^n+b_{n-1}u^{n-1}+\ldots+b_0)T=(uT)^n+b_{n-1}(uT)^{n-1}+\ldots+b_1(uT)+b_0=0.$$

Esta ecuación significa que uT es también raíz de $p(x)$ y, por lo tanto, que uT es un conjugado de u.

Ejemplo 1. Consideremos el campo $K=R(\sqrt{2}, i)$ de grado 4 sobre el campo de los racionales (*) engendrado por $\sqrt{2}$ e $i=\sqrt{-1}$. Sobre el campo intermedio $F=R(i)$, el campo total K es una extensión de grado 2, engendrado por cualquiera de las raíces conjugadas $\pm\sqrt{2}$ de $x^2=2$. Estas dos conjugadas son algebraicamente indistinguibles, lo que es decir (Teor. 7, Cap. XIV) que hay un automorfismo S de K, transformando $\sqrt{2}$ en $-\sqrt{2}$ y conservando fijos

(*) Como en § 5, Cap. XIV, se puede observar que este campo es el campo raíz de x^4-2x^2+9.

los elementos de $R(i)$. El efecto de S sobre cualquier elemento u de K es

$$(2)\qquad (a+b\sqrt{2}+ci+d\sqrt{2}i)S=a-b\sqrt{2}+ci-d\sqrt{2}i.$$

donde hemos representado cada elemento de K mediante los elementos de la base: 1, $\sqrt{2}$, i, $\sqrt{2}i$ (cfr. Cap. XIV, §5). Por un razonamiento similar, existe un automorfismo T, que conserva inalterados los elementos de $R(\sqrt{2})$ y transporta i sobre $-i$. Entonces

$$(3)\qquad (a+b\sqrt{2}+ci+d\sqrt{2}i)T=a+b\sqrt{2}-ci-d\sqrt{2}i,$$

así que, simplemente, T transforma cada número en su complejo conjugado. El producto ST es un tercer automorfismo de K. El efecto de estos automorfismos sobre $\sqrt{2}$ e i se puede tabular como sigue:

$$S\begin{cases}\sqrt{2}\to-\sqrt{2}\\ i\to i\end{cases}\qquad T\begin{cases}\sqrt{2}\to\sqrt{2}\\ i\to -i\end{cases}$$

$$ST\begin{cases}\sqrt{2}\to-\sqrt{2}\\ i\to -i\end{cases}\qquad I\begin{cases}\sqrt{2}\to\sqrt{2}\\ i\to i\end{cases}$$

Decimos ahora que I, S, T y ST son los únicos automorfismos de K sobre R. Por el Teor. 3, cualquier otro automorfismo U debe transformar $\sqrt{2}$ en su conjugado $\pm\sqrt{2}$, e i en su conjugado $\pm i$. Estas son exactamente las cuatro posibilidades tabuladas arriba, I, S, T y ST. Por lo tanto, U debe coincidir con una de ellas en sus efectos sobre los generadores $\sqrt{2}$ e i y, por consiguiente, sobre todo el campo. Así que $U=I$, S, T o ST.

La tabla de multiplicación para estos automorfismos puede construirse directamente partiendo de la tabla ya construída. Resulta:

$$(4)\qquad S^2=I,\qquad T^2=I,\qquad ST=TS.$$

Esta es exactamente análoga a la tabla por los elementos del grupo del rectángulo (Cap. VI, §7), y así concluímos que el grupo de los automorfismos de $R(\sqrt{2}, i)$ es isomorfo al grupo del rectángulo $\{I, S, T, ST\}$.

DEFINICIÓN. *Si* $N=F(u_1, \ldots, u_n)$ *es el campo raíz de un polinomio* $f(x)=(x-u_1)\ldots(x-u_n)$, *el grupo de los automorfismos de*

N sobre F se llama grupo de Galois [sobre F (*)] de la ecuación f(x)=0 o, también, grupo de Galois de N sobre F.

Las propiedades más profundas de las soluciones de una ecuación resultan siempre dependientes de las propiedades de su grupo. En particular, la posibilidad de poder resolver por radicales una ecuación dada, dependerá de que su grupo de Galois tenga ciertas propiedades. Desde este punto de vista, demostraremos que hay ecuaciones de grado superior a cuatro que no pueden resolverse por radicales (ver § 8).

Para describir explícitamente los automorfismos T de un particular grupo de Galois, se procederá como sigue. Sea N el campo raíz de $f(x)$ sobre F. Entonces T representa las raíces de $f(x)$ sobre las raíces de $f(x)$ (Teorema 3), y raíces distintas sobre raíces distintas. Por lo tanto, T efectúa una sustitución ϕ entre las distintas raíces $u_1, \ldots, u_k$, tal como (**)

$$(5) \qquad u_1T=u_{1\phi}, \ \ldots, \ u_kT=u_{k\phi}, \qquad (k \leqslant n).$$

Por otra parte, cada elemento w en el campo de raíces es expresable como un polinomio $w=h(u_1, \ldots, u_k)$ con coeficientes de F. Como T deja inalterados estos coeficientes, la propiedad (5) de T da

$$[h(u_1, \ldots, u_k)]T=h(u_1T, \ldots, u_kT)=h(u_{1\phi}, \ldots, u_{k\phi}).$$

Esta fórmula expresa que el efecto de T sobre w está completamente determinado por el efecto de T sobre las raíces; es decir, que T está unívocamente determinado por la sustitución (5). Como el producto de dos transformaciones se obtiene aplicando sucesivamente los dos automorfismos que les corresponden, las sustituciones (5) forman un grupo isomorfo con el grupo de automorfismos. Las (5) no incluyen necesariamente todas las sustituciones posibles, sino sólo aquellas que conservan todas las identidades polinomiales entre las raíces y pueden así corresponder a automorfismos. Los resultados establecidos se pueden sumariar como sigue:

TEOREMA 4. *Sea f(x) un polinomio de grado n sobre F, el cual tiene exactamente k raíces distintas $u_1, \ldots, u_k$ en un campo raíz*

(*) Cuando no se señale el campo F, se sobreentiende que es el de los coeficientes de $f(x)$. (N. del T.)

(**) Ver N. del T. en Cap. VI, § 7.

$N=F(u_1, \ldots, u_k)$. *En tal caso, cada automorfismo* T *del grupo de Galois* G *de* f(x) *determina una sustitución* $u_i \leftrightarrow u_iT$ *entre las raíces distintas de* f(x), *e inversamente, el automorfismo* T *está completamente determinado por esta sustitución.*

Corolario 1. *El grupo de Galois de un polinomio es isomorfo con un grupo de permutaciones entre sus raíces.*

Corolario 2. *El orden del grupo de Galois de cualquier polinomio de grado* n *es un divisor de* n!

Ejemplo 2. La ecuación $x^4-3=0$ es irreducible sobre el campo R, por el teorema de Eisenstein, y tiene cuatro raíces distintas $r, ir, -r, -ir$, donde $i=\sqrt{-1}$, y $r=\sqrt[4]{3}$ es la raíz cuarta aritmética de 3. El campo raíz $N=R(r, ir, -r, -ir)$ puede ser engendrado como $N=R(r, i)$. Como r es de grado cuatro sobre R, así como i es de grado dos sobre el campo real $R(r)$, el campo de raíces total, N, tiene grado 8 sobre R. Por el Teorema 10, Cap. XIV, esta ampliación N tiene una base de ocho elementos: $1, r, r^2, r^3, i, ir, ir^2, ir^3$. Puesto que cada elemento en N puede ser expresado como una combinación lineal de estos elementos base, con coeficientes racionales, el efecto de un automorfismo T puede ser completamente determinado una vez que rT e iT son conocidos.

Diversos automorfismos de N pueden construirse inmediatamente. Como N es una ampliación de segundo grado sobre el campo real $R(r)$, admite el automorfismo T en el que la imagen de cada número de N es su complejo conjugado; luego $rT=r$, $iT=-i$. Por otra parte, N es una extensión de grado cuatro sobre el subcampo $R(i)$ engendrada por el elemento r. Por el Teor. 7, Capítulo XIV, N admite el automorfismo S que hace corresponder a r su conjugado ir, así que $rS=ir$, $iS=i$. Se deduce que S^2 es un automorfismo con $rS^2=i^2r$, $iS^2=i$, y también, $rS^3=-ir$, $iS^3=i$. Además, por combinaciones de S y T se encuentran para N ocho automorfismos, cuyo efecto sobre los generadores i y r es el siguiente:

	I	S	S^2	S^3	T	TS	TS^2	TS^3
A r le corresponde . .	r	ir	$-r$	$-ir$	r	ir	$-r$	$-ir$
A i le corresponde . .	i	i	i	i	$-i$	$-i$	$-i$	$-i$

Se puede comprobar también que $TS=ST$, $S^4=T^2=I$, así que estos ocho automorfismos forman un grupo. Estos automorfismos constituyen el grupo de Galois completo, pues cualquier automorfismo debe transportar a i sobre uno de sus conjugados $\pm i$, y a r sobre uno de sus conjugados $\pm r$ o $\pm ir$; la tabla anterior incluye, pues, todas las combinaciones posibles de estos efectos.

Muchos conceptos de la teoría de grupos pueden ser aplicados útilmente a los grupos G de Galois. Así, en el ejemplo, G contiene al subgrupo $H=[I, S, S^2, S^3]$ engendrado por S, y al subgrupo $L=[I, S^2]$ engendrado por S^2. Cada automorfismo del subgrupo H deja i invariable y, por lo tanto, deja fijos a todos los elementos del subcampo $R(i)$. El subgrupo inferior L consiste en todos los automorfismos que dejan inalterados todos los elementos del subcampo más amplio $R(i, r^2)$. En este sentido, la sucesión descendente de subgrupos $G \geqslant H \geqslant L \geqslant I$ corresponde a la sucesión ascendente de los subcampos $R \leqslant R(i) \leqslant R(i, \sqrt{3}) \leqslant R(i, r)$. Esta sucesión ascendente de subcampos proporciona un método de resolver la ecuación, adjuntando sucesivamente las raíces de las ecuaciones simples $x^2=-1$, $y^2=3$, $z^2=\sqrt{3}$. Éste es un ejemplo típico de la importancia de los subgrupos del grupo de Galois de una ecuación, para la resolución algebraica de la misma.

En el grupo de Galois aparecen naturalmente los homomorfismos. Así, en el ejemplo anterior, todo automorfismo U del grupo de Galois G transporta a i sobre $\pm i$, luego transportará a cada elemento del campo $R(i)$ sobre algún elemento del mismo campo. Esto significa que U induce un automorfismo U^* de $R(i)$, que será el U^* definido para los elementos w en $R(i)$ por la identidad $wU^*=wU$. La correspondencia $U \to U^*$ es un homomorfismo que representa al grupo G de todos los automorfismos U de N en el grupo G^* de los automorfismos de $R(i)$. Pero G^* consta sólo de dos elementos: la identidad I^* y el automorfismo que intercambia i con $-i$. Además, $U^*=I^*$ si, y sólo si, U deja fijo cada elemento de $R(i)$, esto es, si, y sólo si, U está en el subgrupo $H=[1, S, S^2, S^3]$. Por lo tanto, $U \to U^*$ es el homomorfismo de G en el que al subgrupo H le corresponde la identidad; y el grupo G^* es, por lo tanto, isomorfo con el grupo cociente G/H.

Esta correspondencia entre el subcampo $R(i)$ y el subgrupo H que deja inalterados todos los números del subcampo, es de importancia capital en el desarrollo posterior de la teoría de Galois (§ 5).

EJERCICIOS

1. Dibujar un diagrama para la estructura de red del sistema de todos los subcampos de $R(i, r)$.
2. Demostrar que x^4-3 es irreducible sobre R, viendo que ninguno de los factores lineales o cuadráticos de x^4-3 tiene coeficientes en R.
3. Representar cada automorfismo del grupo de Galois de $x^4-3=0$ como una sustitución entre sus raíces.
4. a) Demostrar que x^4-3 es irreducible sobre $R(i)$.
   b) Hallar el grupo de Galois de x^4-3 sobre $R(i)$.
5. Demostrar, desde los principios, que la siguiente sustitución entre las raíces de x^4-3 no puede corresponder a un automorfismo: $r \to ir$, $ir \to -ir$, $-r \to r$, $-ir \to -r$.
6. Sea $F=R(\omega)$ el campo engendrado por una raíz cúbica ω de la unidad. Discutir el grupo de Galois de x^3-2 sobre F, incluyendo la determinación del grado del campo raíz, una descripción del grupo en lenguaje puramente geométrico y la representación de cada automorfismo como una sustitución.
7. Hacer lo mismo para x^5-7 sobre $R(\zeta)$, siendo ζ una raíz primitiva quinta de la unidad.
8. Hacer lo mismo para x^3-5 sobre R.
9. Demostrar que si ζ es una raíz primitiva n-ésima de la unidad, el grupo de Galois de $R(\zeta)$ es abeliano. (*Sugerencia:* Un automorfismo tiene la forma $\zeta \to \zeta^e$.)
10. a) Si K es una ampliación de R, demostrar que cualquier automorfismo de K deja fijo cada elemento de R.
    b) Enunciar y probar un resultado análogo para los campos de característica p.

3. Polinomios separables e inseparables

La discusión general del grupo de Galois se complica por la presencia de los llamados polinomios irreducibles *inseparables*, los cuales definen elementos algebraicos de grado n que tienen *menos* de n conjugados. Esta complicación acaece en algunos campos de característica p, y puede ser ilustrada con un ejemplo sencillo.

Denotemos por $K=J_p(u)$ una extensión simple trascedente del campo J_p de los enteros módulo p, y sea $F=J_p(u^p)$ el subcampo de K engendrado por $u^p=t$. Así, los elementos de F son todas las formas racionales en un elemento t trascendente sobre J_p. Sobre F, el elemento original u satisface a la ecuación $f(x)=x^p-t=0$. En este caso, el polinomio $f(x)$ resulta irreducible sobre $F=J_p(t)$, pues de no serlo, también sería $f(x)$, por el lema de Gauss (Cap. IV), reducible sobre el dominio $J_p[t]$ de los polinomios en t; pero tal

factorización $f(x)=g(x, t)h(x, t)$ es imposible, ya que x^p-t es lineal en t. Por lo tanto, la raíz u de $f(x)$ tiene grado p sobre F. Pero [cfr. Cap. XIII (36)] $f(x)$ tiene sobre K la factorización

$$(6) \qquad f(x)=x^p-u^p=(x-u)^p.$$

Por lo tanto, tiene *sólo* una raíz, y u (aunque es de grado $p>1$) no tiene conjugados distintos de él mismo.

Podemos describir esta situación con los siguientes términos:

DEFINICIÓN. *Un polinomio* f(x) *de grado* n *se dice separable sobre un campo* F *si tiene* n *raíces distintas en algún campo raíz* $N \geqslant F$; *en caso contrario,* f(x) *se llama inseparable. Una extensión finita* $K \geqslant F$ *se llama separable sobre* F *si cada elemento de* K *satisface sobre* F *a una ecuación polinómica separable.*

Hay un fácil criterio de separabilidad o inseparabilidad, para un polinomio dado $f(x)=a_0+a_1x+\ldots+a_nx^n$. Definamos primero la derivada formal $f'(x)$ de $f(x)$, por la fórmula (cfr. Cap. IV, §1, Ejercicio 7)

$$(7) \qquad f'(x)=a_1+(2\times a_2)x+\ldots+(n\times a_n)x^{n-1},$$

donde $n\times a$ denota el n-ésimo múltiplo natural de a (ver Cap. XIII, §7). Si los coeficientes son del campo de los números reales, esta derivada coincide con la calculada por derivación ordinaria. A partir de la definición (7) se pueden deducir, sin el empleo de los límites, multitud de leyes de derivación, tales como

$$(f+g)'=f'+g', \qquad (fg)'=fg'+gf', \qquad (f^m)'=mf^{m-1}f'.$$

Factoricemos ahora $f(x)$ en potencias de factores lineales distintos sobre su campo raíz N:

$$(8) \qquad f(x)=c(x-u_1)^{e_1}\ldots(x-u_k)^{e_k} \qquad (c\neq 0).$$

Por derivación formal de los dos miembros de (8), hallaremos que $f'(x)$ es la suma de $(k-1)$ términos, cada uno de los cuales contiene a $(x-u_1)^{e_1}$ como factor, y de un término k-ésimo que es $ce_1(x-u_1)^{e_1-1}(x-u_2)^{e_2}\ldots(x-u_k)^{e_k}$. Por lo tanto, si $e_1>1$, $x-u_1$ dividirá a $f'(x)$, y no lo dividirá si $e_1=1$. Repitiendo el razonamiento para $e_2, \ldots, e_k$, resulta que $f(x)$ y $f'(x)$ tienen factores comunes

excepto si $e_1=e_2=\ldots=e_k=1$, es decir, excepto si $f(x)$ es separable; por lo tanto, $f(x)$ factorizado sobre N es separable si, y sólo si, $f(x)$ y su derivada formal $f'(x)$ son polinomios primos relativos.

Pero el m. c. d. de $f(x)$ y $f'(x)$ puede ser calculado como en Capítulo IV, por el algoritmo de Euclides en $F[x]$, y no resultará alterado si F se extiende a un campo más amplio. Por lo tanto,

TEOREMA 5. *Sea* f(x) *un polinomio sobre el campo* F; *sea* d(x) *el m.c.d. (mónico) de* f(x) *y su derivada formal* f'(x), *calculado por el algoritmo de Euclides. Si* d(x)=1, *el polinomio* f(x) *es separable, y en otro caso es inseparable.*

Cuando $f(x)$ sea irreducible, el m. c. d. $[f(x), g(x)]$ es 1, excepto cuando $f(x)$ divide a $g(x)$; y $f(x)$ no puede dividir a un polinomio de menor grado excepto el polinomio nulo. Por lo tanto,

COROLARIO 1. *Si un polinomio irreducible es inseparable, su derivada formal es un polinomio nulo.*

COROLARIO 2. *Cualquier polinomio irreducible sobre un campo de característica* ∞ *es separable.*

Pues $f'(x)=n\times a_n x^{n-1}+\ldots\neq 0$ si $n>0$.

Otra corolario es que si F es de característica ∞, el campo raíz de cualquier polinomio irreducible $f(x)$ de grado n contiene exactamente n raíces conjugadas de $f(x)$. Además, cualquier elemento algebraico sobre un campo de característica ∞ satisface a una ecuación que es irreducible y, por tanto, separable, así que cualquier extensión algebraica de tal campo es separable, en el sentido de la definición precedente.

El resultado del Corolario 2 no es válido para los campos de característica prima. Por ejemplo, el polinomio irreducible x^p-t mencionado al principio de esta sección, tiene una derivada formal $(x^p-t)'=p\times x^{p-1}=0$.

EJERCICIOS

1. Sin emplear el Teor. 5, demostrar que las raíces de un polinomio cuadrático irreducible sobre R son distintas.
2. Sea $f(x)$ un polinomio con coeficientes racionales, y sea $d(x)$ el m. c. d. de $f(x)$ y $f'(x)$. Demostrar que $f(x)/d(x)$ es un polinomio con las mismas raíces que $f(x)$, pero sin raíces múltiples.

3. a) Demostrar que si $f'(x)=0$, $f(x)$ es inseparable sobre cualquier campo F.
   b) Demostrar que si $f'(x)=0$ sobre J_p, entonces $f(x)=[g(x)]^p$ para algún $g(x)$.
4. Demostrar que x^2-2u es inseparable sobre $J_2(u)$. Demostrar que el grupo de Galois de su campo raíz es la identidad.

4. Propiedades del grupo de Galois

Las propiedades generales de los grupos de Galois requieren un análisis más penetrante. En esta sección, enunciaremos primero y demostraremos después los tres teoremas siguientes:

TEOREMA 6. *Dos campos raíces cualesquiera,* N *y* N', *de un polinomio dado sobre* F, f(x), *son isomorfos. El isomorfismo entre* N *y* N' *puede ser elegido de modo tal, que deje fijos los elementos de* F (*).

Este teorema asegura que esencialmente hay un solo campo raíz para un polinomio y, por lo tanto, que los grupos de Galois G y G' obtenidos para dos campos raíces N y N' de $f(x)$ son isomorfos [lo cual se ha admitido tácitamente al decir «el grupo de Galois de $f(x)$»].

TEOREMA 7. *El orden del grupo de Galois de un polinomio separable sobre* F *es exactamente el grado* [N : F] *de su campo raíz.*

En el segundo ejemplo de § 2 hemos visto que éste es efectivamente el caso para el campo raíz de x^4-8.

TEOREMA 8. *En el campo raíz* N $\geqslant$ F *de un polinomio separable, los elementos que permanecen invariables para cualquier automorfismo del grupo de Galois de* N *sobre* F *son, exactamente, los elementos de* F.

Este teorema ofrece alguna información sobre el grupo de Galois G, pues asegura que para cada elemento a en N pero no en F, existe en G algún automorfismo T tal, que $aT \neq a$.

Volvamos ahora al Teor. 6. La afirmación de que el campo raíz es único es, esencialmente, una consecuencia directa del hecho de

(*) Un isomorfismo entre dos campos, que deje fijos los elementos de un subcampo común F, es llamado a veces *equivalencia sobre F*.

que dos raíces diversas de un mismo polinomio irreducible engendran extensiones simples isomorfas (Teor. 7, Cap. XIV). Más concretamente, dos campos raíces dados $N=F(u_1, \ldots, u_n)$ y $N'=F(u_1', \ldots, u_n')$ de un $p(x)$ irreducible comprenden a las extensiones simples $F(u_1)$ y $F(u_1')$ engendradas por las raíces u_1 y u_1' de $p(x)$. Hay, pues, un isomorfismo T de $F(u_1)$ a $F(u_1')$. Falta sólo extender apropiadamente este isomorfismo al campo raíz completo. El procedimiento básico para una tal extensión está dado por

LEMA 1. *Si un isomorfismo* S *entre los campos* F *y* F′ *transforma los coeficientes de un polinomio irreducible* p(x) *en los correspondientes coeficientes de un polinomio* p′(x) *sobre* F′, *y si* F(u) *y* F′(u′) *son extensiones simples engendradas respectivamente por las raíces* u *y* u′ *de estos polinomios, entonces* S *puede ampliarse a un isomorfismo* S* *entre* F(u) *y* F′(u′) *en el cual* uS* = u′.

Demostración. Exactamente como en la discusión del Teor. 7, Capítulo XIV, la extensión deseada S^* está dada explícitamente por la fórmula

$$(9) \quad (a_0+a_1u+\ldots+a_{n-1}u^{n-1})S^*=a_0S+(a_1S)u'+\ldots+(a_{n-1}S)(u')^{n-1},$$

con todos los a_i en F, siendo n el grado de u sobre F.

LEMA 2. *Si un isomorfismo* S *de* F *a* F′ *transforma a* f(x) *en un polinomio* f′(x), *y si* N ⩾ F *y* N′ ⩾ F′ *son, respectivamente, campos raíces de* f(x) *y* f′(x), *el isomorfismo* S *puede ser ampliado a un isomorfismo de* N *a* N′.

Esto será demostrado por inducción relativa al grado $m=[N:F]$. Para $m=1$ es trivial, pues S está «a priori» extendido a N. Sea, pues, $m>1$ y supongamos que el lema es cierto para todos los campos raíces N de grado menor que m sobre algún F. Como $m>1$, no todas las raíces de $f(x)$ están en F, así que hay con seguridad algún factor irreducible de $f(x)$, sea $p(x)$, de grado $d>1$. Sea u una raíz de $p(x)$ en N, mientras que $p'(x)$ sea el factor de $f'(x)$ que corresponde a $p(x)$ en el isomorfismo dado S. El campo raíz N' contiene entonces una raíz u' de $p'(x)$ y, por el Lema 1, el S dado puede extenderse a un isomorfismo S^* con

$$(10) \quad [F(u)]S^*=F'(u'), \quad uS^*=u', \quad p(u)=0, \quad p'(u')=0.$$

Como N está engendrado sobre F por las raíces de $f(x)$, N estará ciertamente engendrado sobre el campo ampliado $F(u)$ por estas mismas raíces, así que N es un campo raíz de $f(x)$ sobre $F(u)$, de grado m/d. Por la misma razón, N' es campo raíz de $f'(x)$ sobre $F'(u')$. Como $m/d < m$, la hipótesis para la inducción de nuestro lema permite afirmar que el isomorfismo S^* de (10) puede ampliarse de $F(u)$ a N. Esto prueba el Lema 2.

LEMA 3. *Si el polinomio* f(x) *del Lema 2 es separable, la ampliación de* S *a* N *podrá hacerse, exactamente, de* m = [N : F] *modos distintos.*

Este resultado puede deducirse por el mismo razonamiento inductivo. Cualquier extensión T del isomorfismo S dado entre F y F' hará corresponder a las raíces u utilizadas en (10) algunas de las raíces u' de $p'(x)$; luego cualquier extensión de S estará producida por una de nuestras construcciones. Como $f(x)$ es separable, su factor $p(x)$ de grado d tiene exactamente d raíces distintas. Esto permite elegir u' de d modos diversos, con lo que tenemos exactamente d posibilidades para la elección de S^* en (10). Por la hipótesis de la inducción, cada una de tales S^* puede ser extendida a N de $m/d = [N : F(u)]$ modos diferentes, luego se tienen en total $d \cdot (m/d) = m$ extensiones, como decíamos.

En el caso de que los dos campos raíces N y N' sean ambas ampliaciones del mismo campo base F, y S sea la identidad que representa a F sobre sí mismo, el Lema 2 demuestra que N es isomorfo con N' y esto demuestra al Teorema 6. Por otra parte, si $f(x)$ es separable y si N y N' son idénticos, el Lema 3 asegura que el automorfismo idéntico I de F puede extenderse de m maneras distintas a un automorfismo de N. Estas ampliaciones son, exactamente, los automorfismos del grupo de Galois de N sobre F, y de aquí el resultado del Teorema 7 sobre el número de automorfismos.

Finalmente, consideremos el Teorema 8; sea G el grupo de Galois del campo raíz N de un polinomio separable sobre F, mientras que K es el conjunto de todos los elementos de N invariantes bajo cualquier automorfismo de G. Se prueba fácilmente que K es un campo y que $K \geqslant F$. Por lo tanto, todo automorfismo en G es una ampliación a N del automorfismo idéntico I de K. Como N es campo raíz de $f(x)$ sobre K, por el Lema 3, hay sólo $[N : K]$ de tales ampliaciones, mientras que, por el Teorema 7, hay $[N : F]$ auto-

morfismos en total. Por lo tanto, $[N:K]=[N:F]$. Como $K \geqslant F$, esto implica que $K=F$, demostrándose así el Teorema 8.

Además, de los lemas sobre la ampliación resulta otra consecuencia, según la cual, el campo raíz es siempre una ampliación normal, en el siguiente sentido:

DEFINICIÓN. *Una ampliación finita* N *de un campo* F *se dice normal sobre* F *si cualquier polinomio* p(x) *irreducible sobre* F *que tenga una raíz en* N *tiene todas sus raíces en* N.

Dicho de otro modo, cualquier polinomio $p(x)$ que sea irreducible sobre F y tenga una raíz en N, puede ser descompuesto en factores lineales sobre N.

TEOREMA 9. *Una ampliación finita de* F *es normal sobre* F *si es campo raíz de algún polinomio sobre* F, *y sólo en este caso.*

Demostración. Si N es normal sobre F, escojamos un elemento u de N pero no de F, y consideremos la ecuación irreducible $p(x)=0$ satisfecha por u. Por la definición de ampliación normal, N contiene todas las raíces de $p(x)$, luego contiene el campo raíz M de $p(x)$. Si hubiese elementos de N no en M, uno de estos elementos v verificaría una ecuación irreducible $q(x)=0$, y M estaría contenido en el campo raíz de $p(x)q(x)$, y así sucesivamente se van formando campos más amplios. Como el grado de N es finito, alguno de los sucesivos campos raíces que así formamos debe ser todo el N.

Ofrece más dificultad mostrar que, recíprocamente, el campo raíz N de cualquier $f(x)$ es normal. Supongamos que hay un polinomio $p(x)$ irreducible sobre F el cual tiene algunas de sus raíces en N, pero no todas. Sea w una raíz de $p(x)$ en N y adjuntemos a N otra raíz w' que no esté en N. La ampliación simple $F(w)$ es isomorfa con $F(w')$ por una correspondencia T en la que $wT=w'$. El campo N es campo raíz de $f(x)$ sobre $F(w)$; por otra parte, $N'=$ $=N(w')$ está engendrado por raíces de $f(x)$ sobre $F(w')$, luego es el campo raíz de $f(x)$ sobre $F(w')$. Por lo tanto, según el Lema 2, la correspondencia T puede extenderse a un isomorfismo de N a N'. Como T conserva invariables los elementos del campo fijado F, estos campos isomorfos N y N' deberán tener el mismo grado sobre F. Pero nosotros hemos supuesto que $N'=N(w')$ es una ampliación propia de N, así que su grado sobre F es mayor que el de N. Esta contradicción demuestra el teorema.

Si la primera parte de esta demostración se aplica a una ampliación separable (en la cual cada elemento satisface a una ecuación separable), todos los polinomios utilizados, $p(x)$, $q(x)$, son separables. Esto demuestra que

COROLARIO. *Cualquier ampliación de* F, *finita, normal y separable, es el campo raíz de un polinomio separable.*

En particular, cualquier ampliación finita y normal N, del campo R de los números racionales, es automáticamente separable (Teorema 5, Corol. 2); luego es el campo raíz de algún polinomio separable. El orden del grupo de automorfismos de N sobre R es, por lo tanto, exactamente, el grado $[N : R]$.

El grupo de Galois puede utilizarse para estudiar las propiedades de los polinomios simétricos (en el sentido del Cap. VI, § 10).

TEOREMA 10. *Sea* $\mathrm{N}=\mathrm{F}(\mathrm{u}_1, \ldots, \mathrm{u}_n)$ *el campo engendrado por las* n *raíces* $\mathrm{u}_1, \ldots, \mathrm{u}_n$ *de un polinomio separable* $\mathrm{f}(\mathrm{x})$ *de grado* n, *y sea* $\mathrm{g}(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ *una forma polinómica sobre* F *simétrica en las* n *indeterminadas* x_i. *El elemento de* N *expresado por* $\mathrm{w}=\mathrm{g}(\mathrm{u}_1, \ldots, \mathrm{u}_n)$ *pertenece precisamente al campo base* F.

Demostración. Cualquier automorfismo T del grupo de Galois G de N efectúa una permutación $u_i \to u_iT$ de las raíces de $f(x)$, por el Teorema 4. La simetría de $g(x_1, \ldots, x_n)$ significa que permanece inalterado por cualquier permutación entre las indeterminadas, luego

$$w \to wT=g(u_1T, \ldots, u_nT)=g(u_1, \ldots, u_n)=w.$$

Como w no es alterada por el automorfismo T, w está en F, por el Teorema 8.

COROLARIO. *Cualquier polinomio (sobre* F*), simétrico en las* n *indeterminadas* $\mathrm{x}_1, \ldots, \mathrm{x}_n$, *puede ser expresado como función racional* (*) *(sobre* F*) de las* n *funciones simétricas elementales*

$$\begin{aligned} \sigma_1&=x_1+x_2+\ldots+x_n, \\ \sigma_2&=x_1x_2+x_1x_3+\ldots+x_{n-1}x_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ \sigma_n&=x_1x_2\ldots x_n. \end{aligned} \tag{11}$$

(*) Y puede ser expresado como polinomio en las σ_i, según puede verse con demostración más detallada.

Para simplificar las fórmulas, expondremos la demostración en el caso $n=3$. Las funciones simétricas elementales σ_1, σ_2, σ_3 engendran sobre F un campo $K=F(\sigma_1, \sigma_2, \sigma_3)$. El campo $N=F(x_1, x_2, x_3)$ engendrado por las tres indeterminadas originales es una extensión finita de K; pues, en efecto, los tres generadores x_i de N son raíces del polinomio cúbico

$$f(x)=(x-x_1)(x-x_2)(x-x_3)=x^3-\sigma_1 x^2+\sigma_2 x-\sigma_3,$$

cuyos coeficientes resultan ser exactamente las funciones simétricas dadas (11). Introduzcamos el grupo de Galois del campo raíz N sobre K. Por el Teorema 4, todo automorfismo de G induce una permutación de las x_i; luego, por el Teorema 10, cualquier polinomio simétrico de las x_i pertenece al campo base K. Como $K=F(\sigma_1, \sigma_2, \sigma_3)$, resulta que tal polinomio simétrico es una función racional de σ_1, σ_2, σ_3.

EJERCICIOS

1. En la demostración del corolario del Teorema 10, mostrar que el grupo de Galois de $N=K(x_1, x_2, x_3)$ sobre K es, precisamente, el grupo simétrico con tres letras.
2. Expresar $x_1^2+x_2^2+x_3^2$ mediante las funciones simétricas elementales.
3. Demostrar que cualquier campo algebraicamente cerrado de característica p, contiene un subcampo isomorfo con el construído en el Apéndice de § 1.
4. a) Demostrar que existen un campo K y un subcampo F tales, que el grupo de Galois de K sobre F es el grupo simétrico de grado n.
   b) Demostrar que, en el Ejercicio 4 a), K puede ser tomado como un subcampo del campo de los números reales. (*Sugerencia:* Utilizar n números reales algebraicamente independientes.)
5. Si un polinomio de grado n tiene n raíces $x_1, \ldots, x_n$, su discriminante es $D=\prod(x_i-x_j)^2$, donde el producto se toma para todos los pares de subíndices con $i<j$.
   a) Demostrar que el discriminante de un polinomio con coeficientes racionales es un número racional.
   b) Para un polinomio cuadrático, expresar D explícitamente como una función racional de los coeficientes.
   * c) La misma cuestión para un polinomio de tercer grado.
6. Demostrar que si K es normal sobre F, y $F \leqslant L \leqslant K$, entonces K es normal sobre L.

5. Subgrupos y subcampos

Si H es un conjunto de automorfismos de un campo N, los elementos a de N que permanecen invariantes para todos los auto-

morfismos de H (de modo que $aT=a$ para todo a en H) forman un *subcampo* de N. En particular, esto será válido si N es el campo raíz de un polinomio sobre un campo base F, y H es un subgrupo del grupo de Galois de N sobre F.

TEOREMA 11. *Si* H *es un grupo finito de automorfismos de un campo* N, *mientras que* K *es el subcampo de todos los elementos invariantes bajo* H, *el grado* [N : K] *de* N *sobre* K *no excede al orden de* H.

Demostración (*). Si H tiene orden n, bastará demostrar que $n+1$ elementos cualesquiera $c_1, \dots, c_{n+1}$ de N son linealmente dependientes sobre K. Con los n elementos T de H construyamos un sistema de n ecuaciones lineales homogéneas

$$y_1(c_1T)+y_2(c_2T)+\dots+y_{n+1}(c_{n+1}T)=0$$

con $n+1$ incógnitas y_i. Tal sistema tiene siempre en N una solución distinta de $y_1=y_2=\dots=y_{n+1}=0$, por el Teorema 10 del Cap. II.

Ahora elijamos el más pequeño entero m para el que las n ecuaciones

$$(12) \qquad y_1(c_1T)+y_2(c_2T)+\dots+y_m(c_mT)=0 \qquad (T \in H)$$

tengan todavía una solución no nula. Esta solución $y_1, \dots, y_m$ consiste en elementos de N y es única, salvo un factor constante, pues si hubiese dos soluciones no proporcionales, una acertada combinación lineal daría una solución del sistema con $m-1$ incógnitas. Sin perjuicio de la generalidad, se puede suponer $y_1=1$. Apliquemos ahora cualquier automorfismo S de H a ambos miembros de (12). Como $TS=T'$ es siempre elemento de H, el resultado es un sistema

$$(y_1S)(c_1T')+(y_2S)(c_2T')+\dots+(y_mS)(c_mT')=0 \qquad (T' \in H)$$

idéntico al (12) excepto el orden de las ecuaciones. Por lo tanto, $y_1S, \dots, y_mS$ es también solución de (12) y, por la unicidad de la solución, debe ser $ty_1, \dots, ty_m$, donde t es un factor de proporcionalidad. Ahora bien, como $y_1=1$ y S es un automorfismo, $y_1S=1$ y también $t=1$. Resulta, pues, que $y_iS=y_i$ para $i=1, \dots, m$, y

(*) Esta demostración, la cual implica la consideración del grupo de Galois, simplemente, como un grupo finito de automorfismos, sin referencia explícita al campo base, se debe al Profesor Artin.

cualquier S en H, lo cual significa que los coeficientes y_i están en el subcampo K de elementos invariantes. La ecuación (12) con $T=I$ muestra entonces que los elementos $c_1, \ldots, c_m$ son linealmente dependientes sobre el campo K. Esto demuestra el teorema.

Basándose en este resultado puede establecerse, al menos para los polinomios separables, una correspondencia entre los subgrupos de un grupo de Galois y los subcampos del correspondiente campo raíz. Esta correspondencia da un método sistemático de reducir las cuestiones relativas a los campos referentes a determinada ecuación, a las cuestiones paralelas referentes a los subgrupos de su grupo (finito) de Galois.

TEOREMA 12 (Teorema fundamental de la Teoría de Galois). *Si* G *es el grupo de Galois para el campo raíz* N *de un polinomio separable sobre* F, *existe una correspondencia biunívoca* H $\leftrightarrow$ K *entre los subgrupos* H *de* G *y aquellos subcampos* K *de* N *que contienen a* F. *Si* K *es dado, el correspondiente subgrupo* H=H(K) *consiste en todos los automorfismos de* G *que conserven fijo cada elemento de* K. *Si* H *es dado, el correspondiente subcampo* K=K(H) *consiste en todos los elementos de* N *invariantes en cualquier automorfismo del subgrupo* H. *Para cada* K, *el subgrupo* H(K) *es el grupo de Galois de* N *sobre* K, *y su orden es el grado* [N : K].

Demostración. Para un K dado, $H(K)$ se define así:

(18) T está en $H(K)$ si, y sólo si, $bT=b$, para todo b en K.

Si S y T tienen esta propiedad, lo mismo sucederá a su producto ST, de modo que $H(K)$ es un subgrupo. El campo N es campo raíz de $f(x)$ sobre K, y cualquier automorfismo de N sobre K es ciertamente un automorfismo de N sobre F, que deja fijo cualquier elemento de K, luego está en el subgrupo $H(K)$. Por lo tanto, $H(K)$ es, por definición, el grupo de Galois de N sobre K. Si se aplica el Teorema 7 a este grupo de Galois, vemos que el orden de $H(K)$ es exactamente el grado de N sobre K.

Dos campos intermedios diversos, K_1 y K_2, determinan distintos subgrupos $H(K_1)$ y $H(K_2)$. Para demostrarlo, elijamos cualquier a en K_1 pero no en K_2, y apliquemos el Teorema 8 al grupo $H(K_2)$ de N sobre K_2. Este teorema afirma que $H(K_2)$ contiene algún automorfismo T en el cual $aT \neq a$. Como a está en K_1, es claro que T no pertenecerá a $H(K_1)$ y, por lo tanto, $H(K_1) \neq H(K_2)$.

Sabemos ahora que $K \to H(K)$ es una correspondencia biunívoca entre *todos* los subcampos de N y *algunos* de los subgrupos de G. Para establecer una correspondencia biunívoca entre *todos* los subcampos y *todos* los subgrupos, deberemos mostrar que cualquier subgrupo aparece como un $H(K)$. Sea H un subgrupo de orden h y definamos $K = K(H)$ como en el enunciado del Teorema 12:

(14) b está en $K(H)$ si, y sólo si, $bS = b$, para todo S en H.

Según el Teorema 11, $[N : K] \leqslant h$. Comparando (13) con (14) se ve que el subgrupo $H(K)$ que corresponde a $K = K(H)$ incluye ciertamente al grupo original dado H, mientras que, por el Teorema 7, el orden de $H(K)$ es $[N : K]$. Como $[N : K] \leqslant h$, esto significa que el orden del grupo $H(K)$ no puede exceder al orden del subgrupo H. Por lo tanto, $H(K) = H$, como enunciábamos. Esto completa la demostración.

El conjunto de todos los campos K entre N y F es una red (Cap. XI, § 8) relativa a la ordinaria relación de inclusión entre subcampos. Si K_1 y K_2 son dos subcampos, su c. i. m. en la red es la intersección $K_1 \frown K_2$, que consiste en todos los elementos comunes a K_1 y K_2, mientras que la c. s. m. es $K_1 \smile K_2$, o sea, el subcampo de N engendrado por todos los elementos en K_1 o K_2; por ejemplo, si $K_1 = F(v_1)$ y $K_2 = F(v_2)$ son dos ampliaciones simples, será $K_1 \smile K_2 = F(v_1, v_2)$.

TEOREMA 13. *La red de todos los subcampos* K_1, K_2, ... *es transportada, por la correspondencia* $K \to H(K)$ *del Teorema 12, sobre la red de todos los subgrupos de* G, *de tal modo que*

(15) $$K_1 \leqslant K_2 \textit{ implica } H(K_1) \geqslant H(K_2),$$

(16) $$H(K_1 \smile K_2) = H(K_1) \frown H(K_2),$$

(17) $$H(K_1 \frown K_2) = H(K_1) \smile H(K_2).$$

En particular, el subgrupo I *corresponde al campo normal completo* N.

El enunciado establece que tal correspondencia invierte la relación de inclusión y transforma la c. i. m. en la c. s. m. (dualidad), e inversamente. Una correspondencia biunívoca entre dos redes que tiene tales propiedades se llama un *isomorfismo dual*.

Para demostrar el Teorema, observemos primero que la definición (13) del grupo relativo a un campo K muestra que cuanto más amplio sea el subcampo, el correspondiente grupo debe dejar más elementos invariables, luego debe ser más restringido. Esto da (15). La c. s. m. y la c. i. m. se definieron con referencia exclusiva a la relación de inclusión (ver Cap. XI); luego, por el principio de dualidad, una correspondencia que invierta la inclusión los intercambia entre sí, como se expresa en (16) y (17).

Omitimos la demostración del siguiente resultado:

TEOREMA 14. *Un campo* K, *con* $N \geqslant K \geqslant F$, *es un campo normal sobre* F *si, y sólo si, el grupo correspondiente* H(K) *es un subgrupo normal del grupo de Galois* G *de* N. *Si* K *es normal, el grupo de Galois de* K *sobre* F *es isomorfo con el grupo cociente* G/H(K).

La tesis de este teorema ha sido ilustrada en un caso particular por el ejemplo al final del § 2.

EJERCICIOS

1. a) Demostrar que si H es un conjunto de automorfismos de un campo N, los elementos de N invariantes para todos los automorfismos de H forman un subcampo K de N.
   b) Demostrar que N es normal sobre este subcampo K.
2. Establecer explícitamente la correspondencia entre subcampos y subgrupos por el campo $R(\sqrt{2}, i)$ sobre R.
3. Lo mismo para el campo de raíces de $x^4 - 3$ estudiado en § 2.
4. Demostrar que el índice de $H(K)$ en G es el grado de K sobre F.
5. Si N es el campo raíz de un polinomio $f(x)$ separable sobre F, demostrar que el número de campos entre N y F es finito.
6. Demostrar que los campos K entre N y F forman una red.
7. Si K es una ampliación finita de un campo F de característica ∞, demostrar que el número de campos entre K y F es finito.
* 8. Demostrar el Teorema 14.
* 9. Dos subcampos K_1 y K_2 en las condiciones del Teorema 12 son llamados conjugados, si existe un automorfismo T de N sobre F en el que K_2 es la imagen de K_1. Demostrar que esto sucede si, y sólo si, $T^{-1}H(K_1)T = H(K_2)$ [esto es, si $H(K_1)$ y $H(K_2)$ son subgrupos «conjugados» de G].

6. Campos finitos

Por el empleo sistemático de las propiedades de los campos raíces, se pueden estudiar de modo completo los campos que constan

de un número finito de elementos (campos finitos) (*). Co campo de característica ∞ contiene siempre un subcampo i isomorfo al de los racionales (Teor. 16, Cap. XIII), cualquie po finito F tiene característica prima p. Sin pérdida de gener podemos suponer que F contiene el campo J_p de los enter dulo p (ver Teor. 13, Cap. XIII). El campo finito F será en una extensión finita de J_p, con una base $u_1, \ldots, u_n$ sobre J_p. quier elemento en F tiene expresión única como combinaci neal $\Sigma a_i u_i$: Cada coeficiente puede tomarse en J_p de p maneras tamente, luego el número total de elementos en F es p^n. Es muestra

Teorema 15. *El número* q *de elementos en un campo es una potencia* p^n *de su característica.*

En un campo finito F con $q=p^n$ elementos, los element nulos forman un grupo multiplicativo de orden $q-1$. El ord cualquier elemento en este grupo es entonces un divisor de así que cualquier elemento satisface a la ecuación $x^{q-1}=1$. tanto, todos los elementos $a_1, \ldots, a_q$ de F (incluyendo al cer tisfacen a la ecuación

(18) $$x^q - x = 0, \quad q = p^n.$$

Por lo tanto, el producto $(x-a_1)(x-a_2)\ldots(x-a_q)$ es visor de x^q-x, ya que es un producto de polinomios primos sí, cada uno de los cuales divide a x^q-x. Como este produc mismo que x^q-x, es mónico y de grado q, obtenemos en clusión

(19) $$x^q - x = (x-a_1)(x-a_2)\ldots(x-a_q).$$

Por lo tanto, F es el campo raíz de x^q-x sobre J_p. Cual otro campo finito F' con el mismo número de elementos es el po raíz de la misma ecuación (**); por lo tanto, es isomorf F, por la unicidad del campo raíz (Teorema 6). Así hemos d trado:

(*) La discusión que sigue depende sólo de las propiedades del campo raí y de su unicidad (Teorema 6), sin que intervengan los grupos de Galois.

(**) O, mejor dicho, de la correspondiente ecuación $1'x^q - 1'x = 0$, donde elemento unidad de F'. El Lema 2 de § 4 permite ampliar el isomorfismo $1 \leftrightarrow 1$ isomorfismo $F \leftrightarrow F'$.

TEOREMA 16. *Dos campos finitos cualesquiera con el misn número de elementos, son isomorfos.*

Consideremos ahora esta cuestión: ¿Cuántos campos finit existen realmente? Para tener un campo finito, formaremos nat ralmente el campo raíz N del polinomio (18) sobre J_p. Este pol nomio $x^q - x$ no tiene factores comunes con su derivada form

$$(x^q - x)' = q \times x^{q-1} - 1 = -1,$$

luego, por el Teorema 5, debe tener q raíces distintas en N. I suma de dos raíces es una raíz, pues $(a \pm b)^p = a^p \pm b^p$ en cualqui campo de característica p, de modo que si $a^{p^n} = a$ y $b^{p^n} = b$, será

$$(a \pm b)^{p^n} = a^{p^n} \pm b^{p^n} = a \pm b.$$

El producto ab es también una raíz, pues $(ab)^{p^n} = a^{p^n} b^{p^n} = a$ y lo análogo vale para el cociente. El conjunto de las raíces c $x^q - x$ es, por lo tanto, un subcampo del campo raíz N, y con este subcampo contiene a todas las raíces, debe coincidir con l totalidad de N. Esto significa que hemos construído un campo cc q elementos y, por lo tanto,

TEOREMA 17. *Siendo* p *un número primo cualquiera y* n *cua quier número natural, existe un campo finito con* $q = p^n$ *elemento éste es el campo raíz de* $x^q = x$ *sobre* J_p.

De los teoremas 16 y 17 resulta que hay esencialmente un s campo con $q = p^n$ elementos. A este campo se le llama a veces *campo de Galois* $GF[p^n]$. La estructura de los grupos multipli tivos de los campos de Galois puede describirse por completo co sigue:

TEOREMA 18. *En cualquier campo finito* F, *el grupo multi cativo de todos los elementos no nulos es cíclico.*

Demostración. Todo elemento no nulo de F es una raíz $(q - 1$ ésima de la unidad, dicho sea en el sentido de que satisface a l ecuación $x^{q-1} = 1$, donde q es el número de elementos de F. Pa probar que el grupo es cíclico se deberá encontrar en F una ra «primitiva» de la unidad de orden $q - 1$, esto es, que no tenga ur potencia igual a 1 con exponente menor que $q - 1$; las potenci

de esta raíz primitiva constituirán entonces la totalidad del grupo A este fin, escribamos $q-1$ como un producto de potencias de factores primos distintos:

$$q-1=p_1^{e_1}p_2^{e_2}\dots p_r^{e_r} \qquad (0<p_1<p_2<\dots<p_r).$$

Para cada $P=p_i$, es $P^e \mid q-1$, así que todas las raíces de $x^{P^e}=1$ son también raíces de $x^{q-1}=1$, luego están todas en F. De estas P^e raíces diversas de la ecuación $x^{P^e}=1$ hay exactamente P^{e-1} que satisfacen la ecuación $x^{P^{e-1}}=1$; por lo tanto, F contiene por lo menos una raíz $c=c_i$ de $x^{P^e}=1$ que no satisface a $x^{P^{e-1}}=1$. Este elemento c_i tiene orden $p_i^{e_i}$ en el grupo multiplicativo de F. El producto $c_1c_2\dots c_r$ será, pues, un elemento de orden $q-1$ (cfr. el siguiente Ejercicio 9), como deseábamos encontrar.

TEOREMA 19. *Cualquier campo finito de característica* p *admite un automorfismo* $a \leftrightarrow a^p$

Demostración. En la discusión general de los campos de característica p, vimos que la correspondencia $a \to a^p$ representa isomorfamente a F sobre el conjunto de las potencias p-ésimas (Capítulo XIII, Teorema 15). Como esta correspondencia es biunívoca, los q elementos a dan exactamente q potencias p-ésimas, las cuales deben, pues, incluir por entero al campo F. Por lo tanto, $a \leftrightarrow a^p$ representa a F sobre la totalidad de F, c. q. d.

COROLARIO. *En un campo finito de característica* p, *cualquier elemento tiene una raíz* p-*ésima.*

Esta conclusión es fundamental en el desarrollo general de la teoría de Galois, pues permite probar que todo polinomio irreducible sobre un campo finito es separable (*), como en el Cor. 2 del Teor. 5 (cfr. § 3, Ejercicio 3 b).

En los ejercicios siguientes se establecerán algunas otras propiedades de los campos finitos.

EJERCICIOS

1. Probar que sobre J_p existen polinomios irreducibles de todos los grados.
2. Probar que cualquier campo finito que contiene a J_p es una ampliación simple de J_p.

(*) Un campo sobre el que cualquier polinomio irreducible sea separable, se llama *perfecto*.

3. Probar que cualquier extensión finita de un campo finito es una extensión simple.
4. a) Utilizando los grados, demostrar que cualquier subcampo de $GF[p^n]$ tiene p^m elementos, donde $m \mid n$.
   b) Si $m \mid n$, demostrar que $(p^m-1) \mid (p^n-1)$.
   c) Empleando b), demostrar que si $m \mid n$, $GF[p^n]$ tiene un subcampo con p^m elementos.
5. Demostrar que la red de todos los subcampos de un campo finito es una cadena.
6. a) En $GF[p^n]$, demostrar que el automorfismo $a \leftrightarrow a^p$ es de orden n.
   b) Demostrar que el grupo de Galois de un campo finito es cíclico.
7. Si m es primo con la característica p de F, demostrar que existe en F una raíz primitiva m-ésima de la unidad. (*Sugerencia:* Aplicar el método usado para Teorema 18. ¿Se puede aplicar esto a un campo de característica ∞?)
8. a) Si $f(x)$ es un polinomio sobre un campo F de característica p con $f'(x)=0$, demostrar que $f(x)$ puede ser escrito en la forma $a_0+a_1x^p+ +\ldots+a_nx^{np}$.
   b) Demostrar que si F es finito, $f(x)=[g(x)]^p$ para un conveniente $g(x)$.
   c) Mediante Ejerc. 8 b), demostrar que todo campo finito es perfecto.
9. Demostrar: en un grupo abeliano, el producto $c_1c_2\ldots c_r$ de los elementos c_i cuyos órdenes son las potencias $p_i^{e_i}$ de distintos números primos, tiene exactamente el orden $p_1^{e_1}\ldots p_r^{e_r}=h$. (*Sugerencia:* Demostrar que el orden divide a h, pero no a h/p_i, para todo i.)
10. a) Mostrar que el grupo multiplicativo de los enteros mód. p (en J_p) es cíclico.
    b) Sea ζ una raíz primitiva p-ésima de la unidad sobre el campo R de los números racionales. Demostrar, mediante a), que el grupo de Galois de $R(\zeta)$ sobre R es cíclico de orden $p-1$.

7. Ecuación cúbica irreducible

La teoría de Galois puede aplicarse a demostrar la imposibiliad de resolver varios problemas clásicos que se relacionan con la ·ón de ecuaciones mediante radicales. Como simple ejemplo e modo de hacer, vamos a considerar el famoso «caso irredu- de la ecuación cúbica con raíces reales.

na ecuación cúbica puede reducirse a la forma

$$f(y)=y^3+py+q=(y-y_1)(y-y_2)(y-y_3) \tag{20}$$

[ver Cap. V, (17)], siendo p y q coeficientes reales, e y_1, y_2, y_3, las tres raíces reales o complejas. Los coeficientes p y q pueden expresarse como funciones simétricas de las raíces, pues efectuando el producto indicado en (20) resulta

$$0=y_1+y_2+y_3, \quad p=y_1y_2+y_1y_3+y_2y_3, \quad q=-y_1y_2y_3. \tag{21}$$

Es importante introducir el discriminante D de la ecuación cúbica, definido por la fórmula

(22) $$D=[(y_1-y_2)(y_1-y_3)(y_2-y_3)]^2.$$

La permutación de dos raíces no altera D; por lo tanto, D es un polinomio simétrico en y_1, y_2, y_3. Por el Teorema 10, resulta que D es expresable como elemento del campo $F=R(p, q)$ de los coeficientes. Tal expresión es, precisamente,

(23) $$D=-4p^3-27q^2.$$

Esta igualdad es una identidad polinómica en y_1, y_2, y_3, como puede evidenciarse atendiendo sucesivamente a las expresiones (21) y (22).

TEOREMA 20. *Una ecuación cúbica real, con discriminante positivo, tiene tres raíces reales; si* D=0, *al menos dos raíces son iguales, mientras que si* D<0, *dos raíces son imaginarias.*

Esto puede comprobarse con la observación de cómo los varios tipos de raíces afectan al valor de D en la fórmula (22). Si todas las raíces son reales, D es, evidentemente, positivo; mientras que $D=0$ si, y sólo si, dos raíces son iguales. Supongamos, finalmente, que una raíz $y_1=a+bi$ es compleja ($b\neq 0$). El complejo conjugado será asimismo raíz, $y_2=a-bi$, mientras que la tercera raíz deberá ser real. En (22), $y_1-y_2=(a+bi)-(a-bi)=2bi$ es imaginario puro, mientras que

$$(y_1-y_3)(y_2-y_3)=(y_1-y_3)(y_1^*-y_3)=(y_1-y_3)(y_1-y_3)^*$$

es un número real. El discriminante D será, por lo tanto, negativo. El Teorema 20 queda así demostrado.

TEOREMA 21. *Si el polinomio cúbico (20) es irreducible sobre* F=R(p, q), *y llamamos* y_1, y_2, y_3 *a sus raíces y* D *a su discriminante, el campo raíz* $F(y_1, y_2, y_3)$ *coincide con* $F(\sqrt{D}, y_1)$.

Demostración. Por la definición (22) de D, el campo raíz contiene ciertamente a $\sqrt{D}$; por lo tanto, falta sólo probar que las raíces y_2 e y_3 están contenidas en $K=F(\sqrt{D}, y_1)$. En este campo,

la cúbica tiene el factor lineal $y-y_1$, así que el restante factor cuadrático

(24) $$(y-y_2)(y-y_3)=y^2-(y_2+y_3)y+y_2y_3$$

tiene también sus coeficientes en K. Sustituyendo en (24) resulta que $(y_1-y_2)(y_1-y_3)$ es de K, de modo que

$$y_2-y_3=\pm\sqrt{D}/(y_1-y_2)(y_1-y_3)$$

también pertenece a K. Pero el coeficiente y_2+y_3 de (24) es asimismo de K. Luego, al pertenecer a K la suma y_2+y_3 y la diferencia y_2-y_3, también pertenecerán y_2 e y_3, c. q. d.

Consideremos ahora una ecuación cúbica que sea irreducible en el campo de sus coeficientes y que tenga sus tres raíces reales. Las fórmulas (19) del Cap. V dan las raíces en la forma $y=z-p/3z$, donde

$$z=\sqrt[3]{-q/2+\sqrt{q^2/4+p^3/27}}=\sqrt[3]{-q/2+\sqrt{-D/108}}$$

[recordando la expresión (23) para D]. Como las raíces son reales, D es positivo (Teorema 20), luego la raíz cuadrada de esta fórmula es un número imaginario. ¡La fórmula obtenida da las raíces *reales* y_i mediante números *complejos*!

Durante mucho tiempo se consideró este hecho como una seria tacha del conjunto de fórmulas resolutivas; los matemáticos se ocuparon largamente intentando expresar las raíces reales de la cúbica por otras fórmulas, en las que sólo interviniesen radicales reales (raíces cuadradas, cúbicas o más elevadas). Estas investigaciones eran en vano, como muestra el siguiente teorema.

TEOREMA 22. *Si un polinomio cúbico tiene sus raíces reales es irreducible sobre el campo* F=R(p, q) *de sus coeficientes, no hay fórmula racional que pueda expresar las raíces de tal ecuación mediante radicales reales sobre* F.

Antes de demostrarlo, discutiremos más por extenso las propiedades de un radical $\sqrt[m]{a}=a^{1/m}$. Si m es compuesto, con $m=rs$, será $a^{1/m}=(a^{1/r})^{1/s}$, es decir, que cualquier radical puede obtenerse mediante una sucesión de radicales con exponente primo. En este último caso, podremos determinar el grado del campo obtenido por adjunción del radical.

LEMA. *Un polinomio* $x^r - a$ *de grado primo* r *sobre un campo real* (*) K, *o es irreducible sobre* K *o tiene una raíz en* K.

Demostración. Adjuntemos a K una raíz primitiva r-ésima de la unidad, ζ, y una raíz u de $x^r - a$. La ampliación resultante $K(\zeta, u)$ contendrá las r raíces, $u, \zeta u, \zeta^2 u, \ldots, \zeta^{r-1}u$, del polinomio $x^r - a$ y, por lo tanto, K es el campo raíz de tal polinomio, el cual admitirá una descomposición factorial

$$(x^r - a) = (x - u)(x - \zeta u)(x - \zeta^2 u) \ldots (x - \zeta^{r-1}u).$$

Supongamos ahora que $x^r - a$ tenga un factor propio e irreducible $g(x)$ de grado $m < r$. Este factor $g(x)$ será el producto de m factores lineales de la descomposición de $x^r - a$ sobre $K(\zeta, u)$, de modo que el término constante b de $g(x)$ será el producto de sus m raíces $\zeta^i u$. Por lo tanto, $b = \zeta^k u^m$ para algún entero k, y

$$b^r = (\zeta^k u^m)^r = (\zeta^r)^k (u^r)^m = (u^r)^m = a^m.$$

Esto nos permitirá hallar en K una raíz r-ésima de a, puesto que al ser $m < r$ y r primo, existirán dos enteros s y t tales, que $sm + tr = 1$ [Cap. I, (11)] y, por lo tanto,

$$b^{sr} = a^{sm} = a^{1-tr} = a/a^{tr},$$

de modo que $a = (b^s a^t)^r$. De este modo, el suponer que $x^r - a$ es reducible sobre K nos conduce a establecer que existe en K una raíz $b^s a^t$ de $x^r - a$, c. q. d.

Volviendo a la demostración del Teorema 22, supongamos que la conclusión sea falsa. Entonces, alguna raíz de la ecuación cúbica podrá ser expresada por radicales reales, lo cual es decir que *alguna* raíz y_1 estará en cierto campo $L = F(\sqrt[r]{a}, \sqrt[s]{b}, \ldots)$ engendrado sobre F por radicales reales. Como D es positivo, al adjuntar el radical real $\sqrt{D}$ se obtiene un nuevo campo real $K = L(\sqrt{D})$. Por el Teorema 21, las raíces de la ecuación cúbica están *todas* en el campo K, así que todas pueden ser expresadas por fórmulas en que intervienen radicales reales. El campo K se obtiene con un número finito de radicales. Si comenzamos por adjuntar $\sqrt{D}$, lo pre-

(*) Un campo real es aquel cuyos elementos son todos reales. Este lema es verdadero para cualquier campo, salvo alguna modificación en la demostración, si K tiene característica r.

cedentemente visto nos enseña que K es el final de una cadena finita de campos

(25) $$F \leqslant K_1 \leqslant K_2 \leqslant K_3 \leqslant \ldots \leqslant K_n = K,$$

donde

(26) $$K_1 = F(\sqrt{D}), \quad K_{i+1} = K_i(a_i^{1/r_i}), \quad i = 1, \ldots, n-1,$$

con cada a_i en K_i y cada r_i primo. Para lograr una ampliación de los campos se puede suponer que a_i^{1/r_i} no pertenece a K_i; esto significa (por el lema), que $x^{r_i} - a_i$ es irreducible sobre K_i y, por lo tanto, que el grado de K_{i+1} es $[K_{i+1} : K_i] = r_i$.

Por la hipótesis, las raíces de la cúbica pertenecen a K; no están en F ni en $F(\sqrt{D})$, porque la cúbica es irreducible sobre F. En la cadena (25) hay, pues, un primer campo K_{j+1} que contiene a una raíz de la cúbica, raíz que designaremos por y_1. Sobre el campo precedente K_j, la cúbica dada debe ser irreducible, pues de otro modo tendría un factor lineal $(y - y_i)$ sobre K_j, contra la hipótesis de que K_j no contiene ninguna y_i. La ampliación

(27) $$K_{j+1} = K_j(a^{1/r}), \quad a = a_j, \quad r = r_j$$

tiene grado r y contiene un elemento y_1 de grado 3 sobre K_j. Por el Teor. 10, Cor. 2, Cap. XIV, es $3 \mid r$, y como r es primo, debe ser igual a 3. En (27) nos encontramos, pues, con una raíz *cúbica* $\sqrt[3]{a}$. El campo K_{j+1} es ampliación del K_j por y_1, contiene a $\sqrt{D}$ y, por el Teorema 21, contendrá a todas las raíces de la cúbica. Por lo tanto, K_{j+1} es el campo raíz de la cúbica dada sobre K_j. Por ser campo raíz será normal, en el sentido del Teorema 9; luego si contiene a una raíz $a^{1/3}$ del polinomio $x^3 - a$ irreducible sobre K_j deberá contener a todas las raíces de este polinomio. Las otras raíces son $\omega a^{1/3}$ y $\omega^2 a^{1/3}$, así que K_{j+1} contiene también a ω, que es una raíz cúbica compleja de la unidad. Esto va contra el supuesto de que $K_{j+1} \leqslant K$ sea un campo real. La demostración queda efectuada.

EJERCICIOS

1. Verificar la fórmula (23) para el discriminante.
2. Mostrar explícitamente las raíces de la cúbica en función de y_1 y $\sqrt{D}$, según el método de Teorema 21.
* 3. ¿Qué parte de la discusión relativa a la ecuación cúbica se puede aplicar a la cúbica sobre un campo de característica p (primo)?
4. Demostrar: un polinomio $x^n - a$ que tiene un factor de grado primo con n, sobre un campo F de característica ∞, tiene una raíz en F.

5. Demostrar: si F es un campo de característica ∞ que contiene a todas las raíces n-ésimas de la unidad, el grado $[F(a^{1/n}):F]$ es un divisor de n.
6. Consideremos el grupo de Galois G de la cúbica (20) irreducible sobre $F=R(p, q)$. Demostrar que si D es el cuadrado de un número de F, G es el grupo alternado de tres letras, y en otro caso es el grupo simétrico.

8. Irresolubilidad de la ecuación de quinto grado

En todo lo que sigue, F denotará un subcampo del campo de los números complejos que contenga a todas las raíces de la unidad, y K denotará una ampliación variable, pero siempre finita, de F.

Supongamos $K=F(a^{1/r})$ engendrado por F y una raíz r-ésima de un elemento $a \in F$, siendo r primo. Las otras raíces de $x^r - a$ son, como en Cap. V, $\zeta a^{1/r}$, $\zeta^2 a^{1/r}$, ..., $\zeta^{r-1} a^{1/r}$, donde ζ es una raíz primitiva r-ésima de la unidad, y, por lo tanto, está en F. Consecuentemente, K es el campo raíz de $x^r - a$ sobre F, y es normal sobre F. Salvo cuando $K=F$, el polinomio $x^r - a$ es irreducible sobre F, por la nota al lema de § 7, luego hay un automorfismo S de K que a la raíz $a^{1/r}$ le hace corresponder la raíz $\zeta a^{1/r}$. Las potencias I, S, S^2, ..., S^{r-1} de este automorfismo hacen que $a^{1/r}$ tenga como imágenes respectivas cada una de las raíces de la ecuación $x^r = a$, con lo que estas potencias incluyen todos los automorfismos de K sobre F. Concluimos, pues, que el *grupo de Galois de K sobre F es cíclico.*

Más generalmente, supongamos que K sea normal sobre F, y que pueda obtenerse por una sucesión de ampliaciones simples sobre F, cada una de las cuales implique solamente la adjunción de una raíz de índice n a la ampliación precedente. Esto quiere decir que existirá una sucesión de campos intermedios K_i,

$$F=K_0 < K_1 < K_2 < \dots < K_s = K, \tag{28}$$

tales que $K_i = K_{i-1}(x_i)$, donde $x_i^{n_i} \in K_{i-1}$. Sin perjuicio de la generalidad, podemos suponer que todos los n_i son primos. Un campo como el K se llama ampliación del F por radicales. Como K es normal sobre F, deberá ser el campo raíz de un polinomio $f(x)$ sobre F y, por ende, el campo raíz del mismo $f(x)$ sobre K_1 y, por lo tanto, normal sobre K_1. Pero K_1 es normal sobre F, por el párrafo precedente. Consecuentemente, todo automorfismo de K sobre F induce un automorfismo de K_1 sobre F, e igualmente la multiplicación de estos automorfismos. Además, por el Lema 2 de

§ 4, cualquier automorfismo de K_1 sobre F puede ampliarse a un automorfismo de K sobre F. Luego la correspondencia inducida es un *homomorfismo* del grupo de Galois de K sobre F al de K_1 sobre F, semejante al descrito al final del § 2. Bajo este homomorfismo, los elementos que inducen el automorfismo idéntico de K_1 sobre F son, precisamente, por definición, los automorfismos de K sobre K_1. Esto demuestra que el grupo de Galois $G(K/F)$ de K sobre F es homomorfo con el $G(K_1/F)$ de K_1 sobre F, siendo este último *isomorfo* con el grupo cociente $G(K/F)/G(K/K_1)$. Combinando esto con el resultado del último párrafo, se infiere que $G(K/K_1)$ es un subgrupo normal de $G(K/F)$, siendo cíclico el grupo cociente $G(K_1/F)$.

Procedamos ahora por inducción sobre s. Por definición, K es una ampliación de K_1 por radicales; como antes, es también normal sobre K_1. Por lo tanto, puede repetirse el precedente razonamiento para $G(K/K_1)$, y demostrar que $G(K/K_2)$ es un subgrupo normal de $G(K/K_1)$, siendo cíclico el grupo cociente $G(K_2/K_1)$. Repitiendo este razonamiento s veces, y designando por S_i el subgrupo $G(K/K_i)$, se llega al siguiente resultado fundamental:

TEOREMA 28. *Sea* K *una ampliación normal de* F *por radicales. En tal caso, el grupo de Galois* G *de* K *sobre* F *contiene una sucesión de subgrupos* $S_0 = G > S_1 > S_2 > \dots > S_s = I$, *cada uno normal en el precedente, siendo cíclicos los grupos cocientes* S_{i-1}/S_i, *y terminando en la identidad* I.

Tal enunciado equivale a decir que el grupo de Galois es *resoluble*, de acuerdo con la siguiente definición:

DEFINICIÓN. *Un grupo finito* G *se llama resoluble cuando contiene una cadena de subgrupos* $S_0 = G \geqslant S_1 \geqslant S_2 \geqslant \dots \geqslant S_s = I$ *tales, que para todo* k, *1.º*, S_k *es normal en* S_{k-1}, *y 2.º*, S_{k-1}/S_k *es cíclico.*

Sobre los grupos abstractos resolubles se conoce mucho; por ejemplo, cualquier grupo cuyo orden es divisible por menos de tres primos distintos es resoluble (Burnside); también es probable que todos los grupos de orden impar sean resolubles. Por ahora, sin embargo, nos contentaremos con el ligero resultado siguiente:

LEMA 1. *Cualquier imagen homomorfa* G' *de un grupo resoluble finito* G *es asimismo resoluble.*

Demostración. Sea G un grupo con la cadena de subgrupos descritos en la definición de resolubilidad, y sean $S_0' = G'$, S_1', ..., $S_n' = I'$ sus imágenes homomorfas. Entonces cada S_k' que contenga a cualesquiera x' e y' también contendrá a $x'y' = (xy)'$ y a $x'^{-1} = (x^{-1})'$ (siendo x e y dos cualesquiera de los elementos de S_k que tienen por correspondientes x' e y' respectivamente) y, por lo tanto. S_k' será un subgrupo de G'. Además, si a está en S_{k-1} y x en S_k, como S_k es normal en S_{k-1}, $a^{-1}xa$ pertenecerá a S_k y, por lo tanto, $a'^{-1}x'a' = (a^{-1}xa)'$ estará en S_k'. Como a' puede ser cualquier elemento de S_{k-1}', queda probado que S_k' es normal en S_{k-1}'. Finalmente, como S_{k-1} consiste en las potencias $(S_k a)^n = S_k a^n$ de alguna clase particular de S_k (siendo S_{k-1}/S_k cíclico), S_{k-1}' consistirá en las potencias $S_k'a'^n = (S_k'a')^n$ de la imagen de esta clase, luego S_{k-1}'/S_k' es también cíclico. La cadena de subgrupos $S_0' \geqslant S_1' \geqslant \ldots \geqslant S_n'$ tiene, por tanto, las propiedades que hacen que G' sea resoluble, como afirma el Lema 1.

Ahora, por definición, diremos que una ecuación $f(x) = 0$, con coeficientes en F, es *resoluble por radicales* sobre F, si sus raíces pertenecen a una ampliación K de F obtenible por sucesivas adjunciones de raíces n-ésimas. Éste es el caso de todas las ecuaciones cuadráticas, cúbicas y cuárticas, por Cap. V, § 5. Se observará que no se requiere que K sea normal, sino sólo que contenga al campo raíz N de $f(x)$ sobre F. Ahora bien, como cualquier conjugado de un elemento expresable por radicales es también expresable por radicales conjugados, el campo N estará también contenido en una ampliación finita $K^* \geqslant K$, que será normal sobre F y también será una ampliación de F por radicales. Entonces K^* contiene a N como un subcampo normal sobre F. Por lo tanto, cada automorfismo de K^* sobre F induce un automorfismo de N sobre F, y la correspondencia es un homomorfismo. Esto es, el grupo de Galois de K^* sobre F es homomorfo con el de N sobre F; pero este último es resoluble, por el Teorema 23; luego, por el Lema 1, también lo es el último. Esto demuestra:

TEOREMA 24. *Si una ecuación* f(x) = 0 *con coeficientes en* F *es resoluble por radicales, su grupo de Galois sobre* F *es resoluble.*

Para demostrar ahora que la ecuación de quinto grado no es siempre resoluble por radicales, nos bastará probar que su grupo

de Galois no siempre es resoluble. Esto lo haremos así: primero probaremos que el grupo simétrico de grado cinco no es resoluble, y después será construída una ecuación de quinto grado cuyo grupo de Galois será el grupo simétrico de grado cinco.

TEOREMA 25. *El grupo simétrico sobre* n *letras no es resoluble, excepto cuando* $n \leqslant 4$.

Demostración. Sea $G = S_0 > S_1 > S_2 > \ldots > S_s$ una cadena de subgrupos, cada uno normal en el precedente y con grupo cociente cíclico S_{k-1}/S_k; vamos a probar, por inducción respecto a s, que S_s *debe contener cualquier ciclo (ijk) de orden 3.* Esto implica que $S_s > I$ y, por lo tanto, G no puede ser resoluble.

Como $S_0 = G$ contiene cualquier ciclo de orden 3, basta probar que si S_{s-1} contiene cualquier ciclo de orden 3, lo mismo le sucede a S_s. Primero, observemos que si las dos permutaciones ϕ y ψ están en S_{s-1}, el producto $\gamma = \phi^{-1}\psi^{-1}\phi\psi$ estará en S_s. Para ver esto, consideremos sus imágenes ϕ' y ψ' en S_{s-1}/S_s. Este último, siendo cíclico, será conmutativo; luego

$$\gamma' = \phi'^{-1}\psi'^{-1}\phi'\psi' = \phi'^{-1}\psi'^{-1}\psi'\phi' = I' \text{ en } S_{s-1}/S_s,$$

lo cual implica $\gamma \in S_s$. Consideremos ahora el caso particular en que $\phi = (ilj)$ y $\psi = (jkm)$, donde i, j, k son dados y l, m son otras dos letras (que existirán si no es $n \leqslant 4$), y tendremos:

$$\gamma = (jli)\,(mkj)\,(ilj)\,(jkm) = (ijk) \in S_s \text{ para todo } i, j, k.$$

Esto demuestra que S_s contiene cualquier ciclo de orden 3, como decíamos. Al producto γ se le llama «conmutador» de los dos ciclos ϕ y ψ.

Incidentalmente, es posible probar una forma más precisa de este teorema. Es sabido que el grupo alternado A es un subgrupo normal del simétrico G, por lo que la cadena comienza $G > A$. Ahora bien, se puede demostrar que el grupo alternado (para $n \geqslant 4$) no tiene subgrupos normales, excepto él mismo y la unidad.

LEMA 2. *Existe (por lo menos) una ecuación real de quinto grado cuyo grupo de Galois es el simétrico con cinco letras.*

Demostración. Sea A el campo de todos los números algebraicos; será numerable y contendrá a todas las raíces de la unidad. Por lo tanto, podremos elegir (ver Cap. XIV, § 6) cinco números

reales $x_1, \ldots, x_5$, algebraicamente independientes sobre A. Formemos la ampliación trascendente $A(x_1, \ldots, x_5)$. Sean $\sigma_1, \ldots, \sigma_5$ las funciones elementales simétricas de las x_i y sea $F = A(\sigma_1, \ldots, \sigma_5)$. Como en el Teorema 10, el grupo de Galois del polinomio

$$f(t) = t^5 - \sigma_1 t^4 + \sigma_2 t^3 - \sigma_3 t^2 + \sigma_4 t - \sigma_5 = 0 \tag{29}$$

sobre F, es el grupo simétrico sobre las cinco letras x_i.

Resulta, pues, por el Lema 2 y el Teorema 25, que existe una ecuación (real) quíntica, sobre un campo que contiene a todas las raíces de la unidad, cuyo grupo de Galois no es resoluble. Aplicando el Teorema 24, tenemos el resultado final.

TEOREMA 26. *La ecuación general de quinto grado no es resoluble por radicales.*

EJERCICIOS

1. Demostrar que el grupo simétrico con tres letras es resoluble.
2. Probar que todo grupo conmutativo finito es resoluble. (*Sugerencia:* Demostrar que contiene un subgrupo normal de índice primo.)
3. Demostrar que si un grupo finito G contiene un subgrupo normal N tal, que N y G/N sean resolubles, G será resoluble.
4. a) Demostrar que en el grupo simétrico con cuatro letras, los conmutadores de ciclos-3 forman un subgrupo normal de orden 4.
   b) Utilizando este resultado y el subgrupo alternado, probar que el grupo simétrico con cuatro letras es resoluble.
5. Demostrar que cualquier grupo abstracto G finito es el grupo de Galois de alguna ecuación. (*Sugerencia:* Por el Teorema de Cayley, G es isomorfo con un subgrupo de un grupo simétrico.)
6. a) Demostrar que el grupo de Galois de $x^n = a$ es también resoluble sobre un campo que no contenga las raíces de la unidad.
   b) Demostrar que el Teorema 24 vale para cualquier F, contenga o no raíces de la unidad.
7. Demostrar detalladamente que si K es una ampliación de F por radicales, existe una ampliación K^* de K que es normal sobre F, y la cual es también ampliación de F por radicales. (Este hecho se ha utilizado en la demostración del Teorema 24.)
8. Si F contiene las raíces n-ésimas de la unidad, y si $K = F(a^{1/n})$, siendo a de K, demostrar que el grupo de Galois de K sobre F es cíclico, aunque n no sea primo.
9. Si R es el campo racional y f el polinomio de (29), demostrar que el grupo de Galois de f sobre el campo $R(\sigma_1, \ldots, \sigma_5)$ es también el grupo simétrico sobre cinco letras.
10. Demostrar que si $n > 4$, existe alguna ecuación de grado n que no es resoluble por radicales.

CAPITULO XVI

Notas ampliatorias

1. Nota al Capítulo V

Ecuaciones de tipo estable. Muchos sistemas físicos son estables cuando, y sólo cuando, todas las raíces de una cierta ecuación polinómica tienen la parte real negativa. Por tal motivo, estas ecuaciones se llaman de «tipo estable» o, simplemente, «estables».

En el caso de la ecuación de segundo grado $z^2+Bz+C=0$ es muy fácil investigar su estabilidad. Si $4C \leqslant B^2$, las dos raíces son reales. Tendrán ambas el mismo signo si, y sólo si, es $z_1 z_2 = C > 0$; tal signo será negativo si, y sólo si, $B = -(z_1+z_2) > 0$. Cuando sea $4C > B^2$, las raíces serán números complejos conjugados. La parte real $x_1 = x_2$ será negativa cuando $B = -2x_1 = -2x_2 > 0$; en este caso también $C > B^2/4 > 0$. Luego, en ambos casos la condición de estabilidad es $B > 0$, $C > 0$.

En el caso de la ecuación cúbica real $z^3+Az^2+Bz+C=0$, tampoco es muy difícil hallar las condiciones de estabilidad. En efecto, si todas las raíces fuesen de parte real negativa, puesto que una raíz $-a$ debe ser real, se tendría la factorización

$$(1) \qquad z^3+Az^2+Bz+C=(z+a)\,(z^2+bz+c),$$

donde $a > 0$ y, por el caso anterior, $b > 0$ y $c > 0$. Por lo tanto, para la estabilidad es necesario que $A = a+b > 0$, $B = (ab+c) > 0$ y $C = ac > 0$. Además, $AB - C = b(a^2+ab+c) > 0$.

Recíprocamente, supongamos $A > 0$, $B > 0$, $C > 0$, y consideremos la factorización real (1), siempre posible. Como $ac = C > 0$, a y c tendrán el mismo signo. Pero si ambos son negativos, b de-

berá ser negativo para que $ab+c>0$, y, por lo tanto, $A=a+b<0$, contra la hipótesis. Por lo tanto, $a>0$ y $c>0$ implican $a^2+ab+c=a(a+b)+c>0$. Pero esto exige que $b=(AB-C)/(a^2+ab+c)>0$ y, por lo tanto, los dos factores de (1) serán estables. Así hemos demostrado el siguiente resultado:

TEOREMA. *La ecuación real de segundo grado* $z^2+Bz+C=0$ *es estable si, y sólo si,* $B>0$ *y* $C>0$. *La ecuación cúbica real* z^3+Az^2+Bz+C *es de tipo estable si, y sólo si,* $A>0$, $B>0$, $C>0$ *y* $AB>C$.

EJERCICIOS

1. Averiguar la estabilidad de las siguientes ecuaciones:
   a) $z^3+z^2+2z+1=0$; b) $z^3+z^2+2z+2=0$.
2. Demostrar que para que un polinomio real mónico de grado n sea estable, todos sus coeficientes deben ser positivos.
3. Demostrar que $z^4+Az^3+Bz^2+Cz+D$ con coeficientes reales es estable si, y sólo si, todos los coeficientes son positivos, $AB>C$ y $ABC>A^2D+C^2$.
4. Teniendo en cuenta el ejercicio anterior, obtener las condiciones necesarias y suficientes para que una ecuación compleja de segundo grado sea estable. [*Sugerencia:* Considerar $(z^2+Bz+C)(z^2+B^*z+C^*)=0$.]

2. Nota al Capítulo VI

Transformaciones «en» y transformaciones «sobre». La noción general de transformación $\phi: S \to T$ de un conjunto S (no vacío) en otro conjunto T, significa una regla ϕ mediante la cual se asigna a cada elemento $p \in S$ un elemento único $p\phi \in T$, que es su imagen. El conjunto S es el *dominio* de ϕ, y T es su *codominio.* Nótese que el conjunto $S\phi$ de las imágenes de elementos en S, esto es, la imagen o resultante de S, puede no agotar el codominio T.

Así la función $f(x)=e^{2\pi ix}$ es una transformación del conjunto R^* de los números reales en el de los números complejos, pero su imagen es, simplemente, la circunferencia unidad.

Cuando el codominio de una transformación es igual a su resultante, así que cualquier $q \in T$ es imagen de al menos un $p \in S$, se llama a ϕ transformación de S sobre T. Cuando por ϕ se transforman elementos distintos en elementos distintos, la transformación de S sobre T se dice *uno-uno* o *biunívoca.*

Dos transformaciones $\phi : S \to T$ y $\phi' : S \to T'$, con el mismo dominio y codominio, se llaman *iguales* cuando dan el mismo resultado sobre cualquier $p \in S$.

(1) $$\phi = \phi' \text{ significa: } p\phi = p\phi' \text{ para todo } p \in S.$$

El *producto* $\phi\psi$ de dos transformaciones ϕ y ψ, supuesto que el dominio de ψ sea el codominio de ϕ, se definirá por la igualdad

(2) $$p(\phi\psi) = (p\phi)\psi.$$

La multiplicación tiene la

Propiedad asociativa: $(\phi\psi)\theta = \phi(\psi\theta)$,

si los productos implicados están definidos. Esto es intuitivamente obvio; formalmente se demuestra por ser

$$p[\phi(\psi\theta)] \underset{\phi(\psi\theta)}{=} (p\phi)(\psi\theta) \underset{\psi\theta}{=} [(p\phi)\psi]\theta \underset{\phi\psi}{=} [p(\phi\psi)]\theta \underset{(\phi\psi)\theta}{=} p[(\phi\psi)\theta]$$

(aplicando la definición de producto (2) al que se indica en cada paso).

La transformación idéntica se definirá por la condición

(3) $$pI = p \quad \text{para todo } p \in S.$$

De las definiciones (1) y (2) resulta fácilmente

Ley de identidad: $I\phi = \phi I = \phi$ para todo ϕ.

En general, si dos transformaciones $\phi : S \to S$ y $\psi : S \to S$ tienen producto $\phi\psi = I : S \to S$, la ϕ se llama inversa a la izquierda de ψ, y la ψ inversa a la derecha de ϕ. Estas definiciones están estrechamente ligadas a los conceptos de «uno-uno» y «sobre» que antes hemos definido.

TEOREMA 1. *La transformación* $\phi : S \to S$ *será «uno-uno» si, y sólo si, tiene inversa a la derecha; y será «sobre» si, y sólo si, tiene inversa a la izquierda.*

Demostración. Si ϕ tiene una ψ con $\phi\psi = I$, la igualdad $p\phi = p'\phi$ implicará

$$p = p(\phi\psi) = (p\phi)\psi = (p'\phi)\psi = p'(\phi\psi) = p',$$

de modo que ϕ será uno-uno. Análogamente, si hay una ψ' tal que $\psi'\phi = I$, cualquier q en S podrá escribirse $q = qI = q(\psi'\phi) = (q\psi')\phi$, y resultará imagen del punto $p = q\psi'$, luego ϕ será sobre.

Viceversa, dada *cualquier* $\phi : S \to S$, construyamos una segunda transformación $\psi : S \to S$ como sigue: para cada q en S que sea imagen de uno o varios p de S, tomaremos como imagen $q\psi$ *uno* cualquiera de estos puntos p. Entonces, para los q de la forma $p'\phi$ resultará $q(\psi\phi) = q$; hagamos que ψ represente todos los puntos q restantes de S sobre cualquier punto fijado del conjunto S.

Con esto, si ϕ es sobre, todo q será de la forma $p\phi$, luego $\psi\phi = I$. Así ψ es inversa a la izquierda de ϕ. Por otra parte, si ϕ es uno-uno, para cada p será $(p\phi)\psi$ el único antecedente p de $q = p\phi$; por lo tanto, $\phi\psi = I$, y ψ será inversa a la derecha de ϕ, como decíamos.

COROLARIO 1. *Una transformación* ϕ : S $\to$ S *es biunívoca de* S *sobre* S *si, y sólo si, tiene inversa por ambos lados, en cuyo caso ambas inversas son iguales.*

En efecto, sean θ y ψ las inversas de ϕ a derecha e izquierda respectivamente. Será

$$\theta = I\theta = (\psi\phi)\theta = \psi(\phi\theta) = \psi I = \psi.$$

La *inversa* de ϕ se define como la transformación ϕ^{-1} que satisface a la

Ley de inversa: $\phi\phi^{-1} = \phi^{-1}\phi = I.$

De lo anterior resulta

COROLARIO 2. *Una transformación* ϕ : S $\to$ S *es una transformación uno-uno de* S *sobre sí mismo si, y sólo si,* ϕ *tiene una inversa* ϕ^{-1}. *En este caso, dos inversas cualesquiera de* ϕ *son iguales, y se tiene*

(4) $$(\phi^{-1})^{-1} = \phi.$$

En el caso especial de ser S finito, la correspondencia será *sobre* si, y sólo si, es *uno-uno*, con lo que huelga la discusión sobre inversas a uno y otro lado.

El teorema y los corolarios valen para transformaciones $\phi : S \to T$, con S y T distintos, bastando observar que las inversas de ϕ son transformaciones del conjunto T en el S

$$\psi\phi = I_T : T \to T, \quad \phi\theta = I_S : S \to S.$$

Aquí I_T e I_S representan, respectivamente, las transformaciones idénticas en S y en T.

Ahora podemos definir un importante concepto : *Grupo de transformaciones* en un «espacio» S, es un conjunto de transformaciones uno-uno de S sobre S tales, que I) la identidad de S pertenece a G ; II) si ϕ está en G, también está ϕ^{-1} ; III) si ϕ y ψ están en G, también está su producto $\phi\psi$.

TEOREMA 2. *El conjunto* G *de todas las transformaciones uno-uno de un espacio* S *sobre sí mismo es un grupo de transformaciones.*

Este es el mismo Teorema 2 del Cap. VI, enunciado ahora en forma más precisa, precisión fácil de trasladar a la demostración allí expuesta.

3. Nota al Capítulo VII

Funciones lineales y espacios duales. Estudiando una ecuación lineal sobre un campo F, $\Sigma c_i x_i = h$, se acostumbra a considerar el primer miembro como una función $f(\xi)$ del vector $\xi = (x_1, \ldots, x_n)$ de $V_n(F)$. De este modo, f indica una transformación que transporta el espacio de las n-plas al campo F de los escalares. Para conformarnos a nuestra notación habitual, haciendo las fórmulas válidas en el caso de que F no sea conmutativo, escribiremos la función f a la derecha de su argumento, poniendo

$$(1) \qquad \xi f = x_1 c_1 + \ldots + x_n c_n, \qquad \xi = (x_1, \ldots, x_n).$$

Esta función es lineal en el sentido de que $(a\xi)f = a(\xi f)$ y $(\xi + \eta)f = \xi f + \eta f$.

Sea V un espacio vectorial sobre un campo F. Llamaremos *función lineal* f sobre V a una función sobre V con valores ξf en F, que verifique las identidades

$$(2) \qquad (\xi + \eta)f = \xi f + \eta f, \qquad (a\xi)f = a(\xi f),$$

para cualesquiera vectores ξ, η de V y cualquier escalar a en F. La primera (2), con $\eta = O$, prueba además que $Of = 0$. Las dos igualdades (2) permiten deducir la siguiente :

$$(3) \qquad (a\xi + b\eta)f = a(\xi f) + b(\eta f), \qquad \xi, \eta \in V\ ;\ a, b \in F.$$

Recíprocamente, de esta primera identidad se deduce la primera de (2), para $a=b=1$, luego $0f=0$, y de aquí, para $b=0$, resulta la segunda de (2). Abreviadamente: una función f es lineal cuando conserva las combinaciones lineales.

Aplicando las precedentes ideas a una combinación lineal de n vectores, se llega a la siguiente caracterización de las funciones lineales de un espacio vectorial.

TEOREMA 1. *Si* $\beta_1, \dots, \beta_n$ *es una base del espacio vectorial* V *sobre* F, *y si* $c_1, \dots, c_n$ *son* n *constantes de* F, *existe una función lineal* f *sobre* V, *y sólo una, tal que* $\beta_i f = c_i$, $i=1, \dots, n$. *Esta función* f *está dada por la fórmula*

$$(4) \qquad (x_1\beta_1+\dots+x_n\beta_n)f=x_1c_1+\dots+x_nc_n.$$

Demostración. Por inducción sobre n, la ecuación (4) se deduce directamente de la (3), para cualquier función lineal f con $\beta_i f=c_i$, $i=1, \dots, n$. Recíprocamente, para toda base $\beta_1, \dots, \beta_n$ de V, cada ξ tiene una expresión única $\xi=x_1\beta_1+\dots+x_n\beta_n$ y, por lo tanto, la expresión (4) define una función uniforme para las constantes $c_1, \dots, c_n$ en F. Esta función es lineal, pues para cualesquiera ξ y $\eta=y_1\beta_1+\dots+y_n\beta_n$ se tiene

$$(a\xi+b\eta)f=[\Sigma(ax_i+by_i)\beta_i]f=\Sigma(ax_i+by_i)c_i=$$
$$=a\Sigma x_ic_i+b\Sigma y_ic_i=a(\xi f)+b(\eta f),$$

cumpliéndose, pues, la condición (3).

COROLARIO. *Las funciones lineales sobre* $V_n(F)$ *son las dadas por las expresiones lineales (1).*

La expresión (1) da, efectivamente, la función f que toma el valor c_i para el vector unidad ϵ_i de $V_n(F)$. Cada función lineal queda así determinada unívocamente por la n-pla $(c_1, \dots, c_n)$ de coeficientes de la fórmula (1); esto sugiere que las funciones lineales forman a su vez un espacio vectorial.

Para cualquier espacio vectorial V podemos definir la suma $f+g$ de dos funciones lineales f y g como la función definida por

$$(5) \qquad \xi(f+g)=\xi f+\xi g, \quad \text{para todo } \xi \text{ en } V.$$

y el producto fc de una función lineal f por un escalar c como una función definida por

$$(6) \qquad \xi(fc)=(\xi f)c, \quad \text{para todo } \xi \text{ en } V,\ c \text{ en } F.$$

Es muy fácil probar directamente que $f+g$ y fc son funciones lineales sobre V.

TEOREMA 2. *Si* V *es un espacio vectorial sobre* F, *el conjunto* V* *de todas las funciones lineales sobre* V *es también un espacio vectorial sobre* F, *definiéndose las operaciones* f+g *y* cf *según las fórmulas (5) y (6).*

El espacio V^* de las funciones lineales sobre V se llama el *espacio dual* o *conjugado* del V. Su consideración es muy importante en el moderno análisis funcional.

La demostración exige, simplemente, verificar que los axiomas que caracterizan a un espacio vectorial son válidos con las operaciones $f+g$ y fc. Por ejemplo, para probar la ley distributiva $(f+g)c=fc+gc$, se observará que, para todo $\xi \in V$,

$$(7) \qquad \xi[(f+g)c]=[\xi(f+g)]c=[\xi f+\xi g]c= \\ =(\xi f)c+(\xi g)c=\xi(fc)+\xi(gc)=\xi(fc+gc),$$

por las definiciones (5) y (6) y la ley distributiva en V. Pero esta igualdad establece que las funciones $(f+g)c$ y $fc+gc$ dan el mismo valor aplicadas a cualquier vector ξ de V, luego son necesariamente la misma función, esto es, son iguales. La demostración de los otros axiomas es análoga.

COROLARIO 1. *Si el espacio vectorial* V *tiene una base finita* $\beta_1, \ldots, \beta_n$, *su espacio dual* V* *tendrá una base* $f_1, \ldots, f_n$ *constituída por las* n *funciones lineales* f_i *definidas por* $(x_1\beta_1+\ldots+x_n\beta_n)f_i=$ $=x_i$, $i=1, \ldots, n$.

Las n funciones lineales f_i quedan determinadas unívocamente por las fórmulas

$$(8) \qquad \beta_i f_j=\begin{cases} 0 \text{ si } i\neq j \\ 1 \text{ si } i=j \end{cases} \qquad i, j=1, \ldots, n.$$

Demostración. Para n escalares cualesquiera $c_1, \ldots, c_n$, la combinación lineal $f=f_1c_1+\ldots+f_nc_n$ es una función lineal; por (8) su valor para cualquier vector β_i de la base es

$$\beta_i(\textstyle\sum_j f_jc_j)=\sum_j\beta_i f_jc_j=c_i.$$

Síguese de aquí que las funciones $f_1, \ldots, f_n$ son linealmente independientes en V^*, ya que si $f=f_1c_1+\ldots+f_nc_n=0$, sería $\beta_i f=0$ para

cada i, luego $c_1=c_2=\ldots=c_n=0$. Y también se deduce que las n funciones $f_1, \ldots, f_n$ describen a V^*, pues, por el Teorema 1, cualquier función lineal f está determinada por los valores $\beta_i f=c_i$ y, por ende, f será igual a la combinación $\Sigma_j f_j c_j$, que los tiene por coeficientes.

La base $f_1, \ldots, f_n$ de V^* es llamada *base dual* de la base $\beta_1, \ldots, \beta_n$ de V.

COROLARIO 2. *El espacio dual* V^* *de un espacio* n-*dimensional* V *tiene también dimensión* n.

La transformación $T: V \to V^*$, que representa a cada vector $\Sigma x_i\beta_i$ de V en la función $\Sigma f_i x_i$ de V^*, es un isomorfismo de V sobre V^*. Este isomorfismo depende de la elección de la base en V.

Cualquier función lineal f sobre $V_n(F)$ queda completamente determinada por los n escalares $c_1, \ldots, c_n$ de la fórmula (1). En la notación general del Cap. VIII, la sucesión de coeficientes $c_1, \ldots, c_n$ se escribe como una matriz C' de tipo $n\times 1$ (columna vector). Así puede considerarse al espacio dual del $V_n(F)$ de filas vectores como el espacio $V_n^*(F)$ de columnas vectores.

Si ξ es un vector de V y f es un vector del espacio dual V^*, puede también escribirse el resultado ξf en notación de «producto interno», $\xi f=(\xi, f)$. En tal caso, la (8) se escribirá

$$(a\xi+b\eta, f)=a(\xi, f)+b(\eta, f), \tag{9}$$

mientras que las definiciones (5) y (6) de adición y multiplicación escalar dan

$$(\xi, fc+gd)=(\xi, f)c+(\xi, g)d. \tag{10}$$

La semejanza de estas igualdades sugiere otra interpretación. En (ξ, f) fijemos ξ y hagamos variar f. Entonces, por (10), ξ determina una función lineal de f y, por (9), las operaciones vectoriales con estas funciones corresponden exactamente a las operaciones vectoriales entre los vectores originales ξ.

Cada ξ en V determina formalmente una función F_ξ en el espacio dual V^*, definida por $F_\xi(f)=(\xi, f)$. Con esto, la fórmula (10) establece que F_ξ es una función lineal.

TEOREMA 3. *Cualquier espacio vectorial* V *de dimensión finita es isomorfo con su espacio conjugado segundo* $(V^*)^*$, *mediante la correspondencia que transporta cada vector* $\xi \in V$ *sobre la función* F_ξ *definida por* $F_\xi(f)=\xi f$.

Demostración. Por (9), la correspondencia $\tau: \xi \to F_\xi$ preserva la adición vectorial y el producto escalar. Veremos ahora que τ es biunívoca y, por ende, un isomorfismo. Si $\xi \neq \eta$, será $\zeta = \xi - \eta \neq O$, y ζ será parte de una base de V. Luego, por el Teorema 1, existirá una función lineal f_0 en V^* con $\zeta f_0 = 1 \neq 0$, tal que

$$F_\xi(f_0) = F_\eta(f_0) + F_\zeta(f_0) = F_\eta(f_0) + 1 \neq F_\eta(f_0)$$

Esto demuestra que la correspondencia τ de V a $(V^*)^*$ es del tipo uno-uno. Pero, por Corol. 2 del Teor. 2, V y $(V^*)^*$ tienen la misma dimensión, luego τ es un isomorfismo de V *sobre* $(V^*)^*$, c. q. d.

Este isomorfismo $\xi \to F_\xi$, a diferencia del que el Corolario 2 establece entre V y V^*, es natural o intrínseco, en el sentido de que su definición no depende de la base elegida en V.

A un subespacio S de V asociaremos el conjunto S' de todas las funciones f lineales en V^* y tales que $(\sigma, f) = 0$ para todo σ en S. Llamaremos a S' el *anulador* de S. Se trata, evidentemente, de un subespacio de V^*, ya que $(\sigma, f) = 0$ y $(\sigma, g) = 0$ implican $(\sigma, fc + gd) = 0$. La correspondencia $S \to S'$ entre los subespacios de V y sus anuladores en V^* tiene la propiedad de que

$$(11) \qquad S \leqslant T \quad \text{implica} \quad S' \geqslant T'$$

(se invierte la inclusión). Pues si $f \in T'$ será $(\sigma, f) = 0$ para cualquier σ en T, luego también para cualquier σ en $S \leqslant T$. El anulador del subespacio constituído por O tan sólo, consiste en todo el espacio dual V^*; y el anulador de todo V es el subespacio de V^* constituído por sólo la función nula.

Dualmente, cada subespacio R del espacio conjugado V^* determina como anulador un subespacio R' de V, constituído por todos los ξ de V con $(\xi, f) = 0$ para cualquier f en R.

TEOREMA 4. *Si* S *es un subespacio* k-*dimensional del espacio* n-*dimensional* V, *el conjunto* S′ *de todas las funciones lineales* f *anulando a* S, *será un subespacio* (n — k)-*dimensional del* V*.

Demostración. Elegida una base $\beta_1, \dots, \beta_k$ de S, se podrá extender a una base $\beta_1, \dots, \beta_n$ de V. En la base dual $f_1, \dots, f_n$ de V^*, la función $f_1c_1 + \dots + f_nc_n$ se anulará en todo S si, y sólo si, se anula para cada $\beta_1, \dots, \beta_k$, o sea, si $c_1 = \dots = c_k = 0$. Esto significa, precisamente, que las $n - k$ funciones $f_{k+1}, \dots, f_n$ constituyen una base del anulador S' de S.

Este Teorema permite formular de otro modo el Teorema r
tivo al número de soluciones independientes de un sistema li
y homogéneo de ecuaciones.

La correspondencia $S \to S'$ entre los espacios y sus anulad
lleva al Principio de Dualidad de la geometría proyectiva n-dim
sional, en la cual son también básicos los resultados que sigu

TEOREMA 5. *La correspondencia* $S \to S'$ *cumple las siguie propiedades:*

(12) $$(S')' = S, \quad (S+T)' = S' \frown T', \quad (S \frown T)' = S' + T'.$$

Demostración. Puesto que $(\xi, f) = 0$ para todo ξ en S y t
f en S', cada ξ en S anulará a cualquier vector f en S', luego $\xi \in ($
o sea $(S')' \geqslant S$. Pero por el Teorema 4, la dimensión de (S')
$n - (n - k) = k = d[S]$; por lo tanto es imposible que $(S')' > S$
será $(S')' = S$.

Esta ecuación establece que la correspondencia $S \to S'$ de
subespacio a su anulador aplicada dos veces es la identidad, y,
tener inversa, esta correspondencia será biunívoca de S *sobre*
Por esto, también puede invertirse la inclusión expresada en (1
y de aquí que $S + T$, el menor espacio que contiene a S y a T
transformará en $S' \frown T'$, el mayor espacio contenido en S' y en
Dualmente, $(S \frown T)' = S' + T'$.

COROLARIO 1. *Sea* L(V) *el conjunto de todos los subespa de un espacio vectorial* V *de dimensión finita sobre un campo c mutativo. Existe una correspondencia biunívoca de* L(V) *sobre mismo, que invierte la inclusión y satisface a (12).*

Demostración. Fijemos en V una base $\beta_1, \ldots, \beta_n$ cualqui
Sea S un subespacio cualquiera de V, y sea S' el conjunto de to
los vectores $\eta = y_1\beta_1 + \ldots + y_n\beta_n$ tales que

(13) $$x_1 y_1 + \ldots + x_n y_n = 0 \text{ para todo } \xi = (x_1\beta_1 + \ldots + x_n\beta_n) \in S.$$

Puede repetirse ahora el razonamiento que conduce a (12) y al T
rema 4, lo que nos llevará al resultado deseado.

Nota 1. Supongamos que V es un espacio vectorial sobre
campo D no necesariamente conmutativo (*). Consideremos V co

(*) A los campos D conmutativos o no, se les llama también *anillos de divi* (division ring), reservando la palabra «campo» para el caso conmutativo.

un espacio vectorial *izquierdo* (escalares a la izquierda). Las fórm las anteriores, especialmente la (7), indicarán que el dual V^* es espacio vectorial *derecho* (los múltiplos escalares a la derecha de vectores). Con esta convención pueden extenderse casi todos los sultados de esta sección, aunque no el anterior Corolario 1 ni el i morfismo no intrínseco del espacio V con su dual V^*.

Nota 2. En el caso de un espacio E vectorial euclídeo. exi un isomorfismo natural, o intrínseco, entre E y su dual E^*, q puede definirse mediante el producto interno (ξ, η). La fórm $\xi f_\eta = (\xi, \eta)$ define para cada vector $\eta \in E$ una función f_η sobre E, q es lineal, puesto que (ξ, η) es bilineal. Puede verse fácilme que la correspondencia $\eta \to f_\eta$ es un isomorfismo de E sobre E^*.

Nota 3. El isomorfismo entre V y V^* no será generalme válido para un espacio V de infinitas dimensiones. Sea V, por eje plo, el conjunto de todas las sucesiones $\xi = (x_1, \ldots, x_n)$, $x_n \in F$, q tengan un número finito de elementos distintos de cero. La adic y multiplicación se efectuarán término a término. Cualquier fu ción lineal sobre V puede ser representada en la forma $\xi f = \Sigma$ por una lista infinita arbitraria de coeficientes $(c_1, c_2, \ldots, c_n,$. Luego el espacio dual V^* constará de tales listas o sucesiones i nitas. Los espacios V y V^* no son isomorfos. Por ejemplo, si F un campo numerable, V será numerable, pero V^* no lo será.

EJERCICIOS

1. Completar la demostración del Teorema 2.
2. Sean n funciones $f_1, \ldots, f_n$ linealmente independientes sobre un esp n-dimensional V, y sean $c_1, \ldots, c_n$ constantes dadas. Probar que ex un vector ξ en V, y sólo uno, con $\xi f_i = c_i$, $i = 1, \ldots, n$. Interpretarlo diante las ecuaciones lineales no homogéneas.
3. a) Completar la prueba de la Nota 2.
   b) Mostrar su relación con el Corol. 1 del Teor. 1.
4. En $V_4(C)$ definir $(\xi, \eta) = x_1 y_2 - x_2 y_1 + x_3 y_4 - y_3 x_4$. Para cada subespac definir S' como el conjunto de todos los vectores η con $(\xi, \eta) = 0$ para $\xi \in S$. Demostrar (11)-(12) y ver que si S es unidimensional, será S <

4. Nota al Capítulo IX

Geometría proyectiva. En el plano real afín, dos puntos tintos pertenecen a una sola recta, y dos rectas *no paralelas* se «

tan en un punto único. Construyamos ahora un plano real *proyectivo*, en el cual

I) Dos puntos distintos pertenecen a una recta única ;
II) Dos rectas distintas *cualesquiera* se cortan en un punto único.

Estas dos propiedades de incidencia son duales entre sí, en el sentido de que cambiando la palabra «recta» por «punto» y viceversa, la I) se cambia en la II) y la II) en I).

Una manera de construir el plano proyectivo real $P_2(R^*)$ es la siguiente : Sea V_3 un espacio vectorial tridimensional sobre el campo R^* de los números reales. Llamemos «punto» de P_2 a un subespacio vectorial lineal y unidimensional de V_3, tal como S. Llamaremos «recta» de P_2 a un subespacio vectorial lineal y de dimensión 2 de V_3, tal como L. Finalmente, diremos que S está en L, o pertenece a L, cuando el subespacio S esté contenido en el subespacio L.

La demostración de que los «puntos» y «rectas» de $P_2(R^*)$ cumplen I) y II) es como sigue : Si los puntos S_1 y S_2 son los espacios vectoriales desarrollados por los vectores α_1 y α_2, será $S_1 \neq S_2$ cuando, y sólo cuando, α_1 y α_2 sean linealmente independientes. La única recta L en que están S_1 y S_2 será entonces el subespacio vectorial bidimensiosal desarrollado por α_1 y α_2, lo que demuestra I). En segundo lugar, si dos rectas (subespacios vectoriales bidimensionales) L_1 y L_2 son distintas, el subespacio $L_1 + L_2$, que es su suma lineal, deberá tener una dimensión mayor, luego coincidirá con todo V_3. Por lo tanto,

$$\dim (L_1 \frown L_2) = \dim L_1 + \dim L_2 - \dim (L_1 + L_2) = 2+2-3=1,$$

de modo que el único punto común a L_1 y L_2 será el espacio unidimensional $L_1 \frown L_2$, lo que demuestra II).

Para obtener en $P_2(R^*)$ unas coordenadas proyectivas convenientes, supongamos que V_3 es el espacio $V_3(R^*)$ de las ternas (x_1, x_2, x_3) de números reales. Entonces, cada terna no nula determinará un punto S de P_2. Las dos ternas $(x_1, x_2\ x_3)$ y (cx_1, cx_2, cx_3) determinarán el mismo punto si $c \neq 0$. A estas ternas, con la identificación

$$(x_1, x_2, x_3) = (cx_1, cx_2, cx_3), \quad c \neq 0,$$

las llamaremos *coordenadas homogéneas* del punto S. Y puesto que un subespacio bidimensional L de V_3 puede definirse como el conjunto de vectores solución de una ecuación lineal homogénea, una recta L de P_2 será el conjunto de puntos cuyas coordenadas homogéneas satisfacen a una ecuación

$$(1) \qquad a_1x_1+a_2x_2+a_3x_3=0 \qquad (a_1, a_2, a_3)\neq(0, 0, 0).$$

Llamaremos a (a_1, a_2, a_3) las *coordenadas homogéneas de la recta L*. Es claro que las ternas (a_1, a_2, a_3) y (ca_1, ca_2, ca_3) determinan la misma recta.

El plano proyectivo real tiene una representación geométrica muy simple. Las coordenadas homogéneas de un punto pueden ser «normalizadas» multiplicándolas por $(x_1^2+x_2^2+x_3^2)^{-1/2}$, así que en las nuevas coordenadas (y_1, y_2, y_3) se tendrá $y_1^2+y_2^2+y_3^2=1$, pudiéndose representar sobre la esfera unidad. Dos puntos diametralmente opuestos de esta esfera, (y_1, y_2, y_3) y $(-y_1, -y_2, -y_3)$, determinan el mismo punto de P_2. De otro modo: los puntos de P_2 pueden obtenerse identificando los pares de puntos diametralmente opuestos de la esfera unidad. Ahora bien, los subespacios bidimensionales L de V_3 cortan a esta esfera en círculos máximos, de modo que cada recta de P_2 consistirá en los puntos (o pares de puntos antípodas) de un círculo máximo. Así, es claro que dos rectas proyectivas (dos círculos máximos) se cortan siempre en un punto (un par de puntos antípodas).

Es claro también que cada subespacio vectorial unidimensional (cx_1, cx_2, cx_3), con $x_3\neq0$, corta al plano afín $x_3=1$ en el punto $(x_1/x_3, x_2/x_3, 1)$; las razones $(x_1/x_3, x_2/x_3)$ se llaman *coordenadas no homogéneas* del punto proyectivo (cx_1, cx_2, cx_3). Pero el lugar $x_3=0$ es una recta proyectiva, llamada *recta del infinito*. Se comprueba que toda recta

$$L: \quad a_1x_1+a_2x_2+a_3x_3=0$$

del plano proyectivo P_2 es, o bien la recta del infinito $x_3=0$ (si $a_1=a_2=0$), o bien una recta del plano afín, $a_1(x_1/x_3)+a_2(x_2/x_3)+$ $+a_3=0$, más un punto $(a_2, -a_1, 0)$ de la recta del infinito.

Sobre cualquier campo F puede construirse un espacio proyectivo n-dimensional. El paso esencial comienza en la consideración de un espacio vectorial $V=V_{n+1}(F)$ de una dimensión más. Luego, $P=P_n(F)$ se construye como sigue: un *punto* de P es un subespa-

cio S unidimensional de V; un subespacio *m-dimensional* de l el conjunto de todos los puntos S de P que pertenecen a un s espacio $(m+1)$-dimensional de V, tal como L. Evidentemente, c subespacio es isomorfo al espacio proyectivo m-dimensional P_m terminado del mismo modo por el espacio vectorial $(m+1)$-din sional L. Si representamos V (por coordenadas con respecto a base conveniente) como el espacio constituído por conjuntos $(n+1)$ elementos ordenados de F, entonces, cada punto S de vendrá dado por sus $n+1$ coordenadas homogéneas $(x_1, \ldots, x_{n+1})$ y las coordenadas $(cx_1, \ldots, cx_{n+1})$ con $c \neq 0$, determinarán el r mo punto. Un hiperplano (o subespacio de dimensión $n-1$) s de nuevo el lugar definido por una ecuación lineal homogénea

$$(2) \qquad a_1x_1 + \ldots + a_{n+1}x_{n+1} = 0, \quad (a_1, \ldots, a_{n+1}) \neq (0, \ldots, 0).$$

Los números $(a_1, \ldots, a_{n+1})$ pueden considerarse como las coorde das homogéneas del hiperplano. Las relaciones entre el espacio p yectivo P y el «espacio proyectivo» cuyos puntos son los hiperp nos de P, son exactamente paralelas a las relaciones entre el espa vectorial V y el espacio dual V^*. Por los resultados del Cap. V sabemos, además, que el conjunto de r ecuaciones del tipo (2) nealmente independientes determinará un subespacio proyectivo $n-r$ dimensiones.

Sea $T: V \to V$ una transformación lineal regular. Es sabido q T transformará cada subespacio lineal S de V en un subespa lineal S^* de V. Por lo tanto, T induce una transformación $S \to S^*$ de los puntos del espacio proyectivo P, y esta transformación hace corresponder a subespacios proyectivos también subespac proyectivos, conservando su dimensión. Llamaremos a T^* u *transformación proyectiva* de P. Si T_1 y T_2 son dos transformac nes regulares de V, el producto T_1T_2 inducirá una transformac $(T_1T_2)^*$ de P, igual al producto $T_1^*T_2^*$ de las transformaciones ducidas. Por lo tanto, el conjunto de todas las transformaciones p yectivas constituye un grupo, que es el *grupo proyectivo n-dim sional*. La correspondencia $T \to T^*$ será un homomorfismo del gru lineal completo en $n+1$ dimensiones al grupo proyectivo de n mensiones.

Con referencia a un sistema dado de coordenadas de V, la tra formación T vendrá dada por una matriz regular $\|a_{ij}\|$ de t $(n+1) \times (n+1)$. En este caso, la transformación T^* transform

el punto de coordenadas homogéneas $(x_1, \ldots, x_{n+1})$ en el punto d coordenadas homogéneas $(y_1, \ldots, y_{n+1})$ dadas por

$$(3) \qquad y_j = x_1 a_{1j} + \ldots + x_{n+1} a_{n+1,j} \quad (j=1, \ldots, n+1).$$

TEOREMA 1. *Las matrices* $(n+1)\times(n+1)$, A, *determinan* l *transformación proyectiva idéntica* T^* *de* P_n *si, y sólo si, A es u múltiplo escalar* cI *de la matriz identidad* I, *con* $c \neq 0$.

Demostración. Si $A = cI$, será $y_j = cx_j$ y las coordenadas homo géneas $(x_1, \ldots, x_{n+1})$ y $(cx_1, \ldots, cx_{n+1})$ determinarán el mismo pun to, luego T^* será la identidad. Recíprocamente, supongamos qu T^* es la identidad. En este caso, cada uno de los $n+1$ vectore unidad ϵ_i deberá transformarse en un múltiplo escalar $c_i\epsilon_i$, lueg A será una matriz diagonal con elementos diagonales $c_1, \ldots, c_{n+1}$ Pero también el vector (1, 1, ..., 1) deberá transformarse en un de sus múltiplos escalares, y como A lo transforma en $(c_1, \ldots, c_{n+1})$ deberán ser iguales todas las c_i y, por consiguiente, A será un múl tiplo escalar de I, c.q.d.

COROLARIO. *El grupo proyectivo en* n *dimensiones sobre* e *campo* F, *es isomorfo con el grupo cociente del grupo lineal com pleto en* n+1 *dimensiones por el subgrupo de los múltiplos esca lares no nulos de la identidad.*

Demostración. La representación $T \to T^*$ es un homomorfis mo del grupo lineal completo al grupo proyectivo. El Teorema dice que el núcleo de este homomorfismo es, precisamente, el con junto de múltiplos escalares de la identidad. Y de aquí resulta l tesis, recordando el Teor. 26 del Cap. VI.

También es consecuencia de esto que dos matrices A y A_1 de terminarán la misma transformación si, y sólo si, es $A_1 = cA$, co c escalar.

Para una recta proyectiva unidimensional, la transformació proyectiva tiene la forma

$$(4) \qquad \begin{aligned} y_1 &= ax_1 + bx_2 \\ y_2 &= cx_1 + dx_2 \end{aligned} \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0.$$

En función de las coordenadas no homogéneas, $z = x_1/x_2$ y $w = y_1/y_2$ esta transformación puede expresarse por la sustitución fracciona ria de términos lineales

$$(5) \qquad w = \frac{cz+d}{az+b},$$

que se obtiene dividiendo miembro a miembro las (4). Esta e
ción puede interpretarse así: cuando $c=0$, el punto $z=\infty$ se tr
forma en el $w=\infty$; si $c\neq 0$, el punto $z=\infty$ se transforma en el
mientras el punto $z=-d/c$ se transforma en el $w=\infty$. La cor
ción de esta interpretación simbólica puede establecerse volvie
a las ecuaciones homogéneas (4). En el espacio de n dimensi
es posible una representación semejante de las transformaci
proyectivas, poniéndose

$$(5') \qquad w_i = \frac{z_1a_{1i}+\ldots+z_na_{ni}+a_{n+1,i}}{z_1b_1+\ldots+z_nb_n+b_{n+1}} \qquad (b_j=a_{j,n+1},\ i=1,\ \ldots,\ n).$$

Hemos visto que las transformaciones proyectivas de P_1
transforman rectas en rectas. Recíprocamente, es un resultado
sico que toda transformación biunívoca de un espacio proyec
real, que transforme rectas en rectas, es proyectiva, para n
(ver Ejercicio 6).

Una forma cuadrática homogénea con tres variables determ
un lugar

$$(6) \qquad \sum_{i,j} x_ib_{ij}x_j=0 \qquad [i,\ j=1,\ 2,\ 3]$$

en el plano proyectivo, pues si las coordenadas $(x_1,\ x_2,\ x_3)$ sati
cen a la ecuación, lo mismo sucederá con las $(cx_1,\ cx_2,\ cx_3)$. A
lugar se le llama una *cónica proyectiva;* la característica (pro
tiva) de la cónica es igual a la característica de la matriz B de
coeficientes. En el plano proyectivo real, cualquier cónica no de
nerada es equivalente a una de las que tienen por ecuación alg
de las siguientes:

$$(7) \qquad x_1^2+x_2^2+x_3^2=0, \qquad x_1^2+x_2^2-x_3^2=0,$$

$$(7') \qquad -x_1^2-x_2^2-x_3^2=0, \qquad x_1^2-x_2^2-x_3^2=0.$$

Cambiando todos los signos no altera el lugar representado, lu
las cónicas (7') son las mismas (7). Pero la primera de ellas car
de puntos reales. Por lo tanto, puede concluirse que *en el pl
proyectivo real, dos cónicas cualesquiera son proyectivamente e
valentes.*

EJERCICIOS

1. En el espacio proyectivo tridimensional sobre un campo F, demostr
   a) Dos puntos distintos determinan una recta;
   b) Tres puntos distintos no alineados determinan un plano.

2. Generalizar el Ejercicio 1 al espacio proyectivo n-dimensional.
3. Expresar todos los puntos, las rectas y los puntos de cada recta en plano proyectivo sobre J_3.
4. En el plano proyectivo sobre un campo finito de n elementos, demostr que existen n^2+n+1 puntos, n^2+n+1 rectas y $n+1$ puntos sobre ca recta.
5. La razón doble de cuatro puntos z_1, z_2, z_3 y z_4 es, por definición, el c ciente $(z_3-z_1)(z_4-z_2)/(z_3-z_2)(z_4-z_1)$ (con el oportuno convenio cua do es $z_1=\infty$). Demostrar que la razón doble es invariante para las trar formaciones lineales (5).
6. Demostrar que la transformación $(z_1, z_2, z_3) \to (z_1^*, z_2^*, z_3^*)$ transpor rectas sobre rectas, en el plano proyectivo complejo, pero no es proye tiva. El asterisco indica al complejo conjugado.

¿Qué representa la cónica proyectiva $x_1^2=2x_2x_3$ en el plano afín?

Dar una clasificación completa respecto al grupo proyectivo de las cu dricas en $P_3(R^4)$.

BIBLIOGRAFIA

TRATADOS GENERALES

ALBERT, A. A.: *Modern Higher Algebra*. Chicago, The University of Chicag Press, 1937.

BÔCHER, M.: *Introduction to Higher Algebra*. New York, Macmillan, 1907.

BOURBAKI, N.: *Éléments de Mathématique*. Livre II, Algèbre. Paris, Herman et Cie. Chap. I, Structures Algèbriques, 1942; Chap. II, Algèbre Linéair 1947; Chap. III, Algèbre Multilinéaire, 1948; Chap. IV, *Polynomes*, 1950 Chap. V, Corps Conmutatifs, 1950.

DICKSON, L. E.: *Modern Algebraic Theories*. Chicago, B. H. Sanborn an Co., 1926.

DUBREIL, P.: *Algèbre*, Paris, Gauthier-Villars. Vol. I. 1946.

JACOBSON, N.: *Lectures in Abstract Algebra*. New York, D. Van Nostrand Cc Inc., Vol. I, 1951; Vol. II, 1952.

MACDUFFEE, C. C.: *An Introduction to Abstract Algebra*. New York, Joh Wiley and Sons, 1940.

PICKERT, G.: *Einführung in die höhere Algebra*. Göttingen, Vandenhoec und Ruprecht, 1951.

SCHREIER, O., and SPERNER, E.: *Introduction to Modern Algebra and Matri Theory*. New York, Chelsea Publishing Co., 1952.

THOMAS, J. M.: *Theory of Equations*. New York, MacGraw-Hill, 1938.

USPENSKY, J. V.: *Theory of Equations*. New York, McGraw-Hill, 1948.

VAN DER WAERDEN, B. L.: *Modern Algebra*. Traducción inglesa: New Yor Frederick Ungar Publishing Co.; Vol. I, 1949; Vol. II, 1950.

WEISNER, L.: *Theory of Equations*. New York, Macmillan, 1938.

TEORIA DE NÚMEROS

DICKSON, L. E.: *Modern Elementary Theory of Numbers*. Chicago, The Ur versity of Chicago Press, 1939.

HARDY, G. H., y WRIGTT, E. M.: *An Introduction to the Theory of Numbe* Oxford, Clarendon Press, 1945.

NAGELL, T.: *Introduction to Number Theory*. New York, John Wiley a Sons, 1951.

ORE, O.: *Number Theory and Its History*. New York, McGraw-Hill, 1949.

STEWART, B. M.: *Theory of Numbers*. New York, MacMillan, 1952.

USPENSKY, J. V., y HEASLET, M. H.: *Elementary Number Theory*. New Yo McGraw-Hill, 1939.

TEORIA DE NÚMEROS ALGEBRAICOS

HASSE, H.: *Zahlentheorie*. Berlin, Akademie-Verlag, 1949.

HECKE, E.: *Vorlesungen über die Theorie der algebraischen Zahlen*. Leipz Akademische Verlagsgesellschaft, 1923. También en New York, Chel Publishing Co., 1948.

POLLARD, H.: *The Theory of Algebraic Numbers*. Carus Mathematical Mo graph N.º 9, New York, The Mathematical Association of America and Jo Wiley and Sons, 1950.

WEYL, H.: *Algebraic Theory of Numbers*. Princeton, Princeton Univers Press, 1940.

TEORIA DE GRUPOS

CARMICHAEL, R. D.: *Introduction to the Theory of Groups of Finite Ord* Boston, Ginn and Company, 1937.

SPEISER, A.: *Theorie der Gruppen von endlicher Ordnung*. 3.ª edic., B Springer, 1937.
ZASSENHAUS, H.: *The Theory of Groups*. Traducción del alemán, por Kravtez. New York, Chelsea Publishing Co., 1949.

TEORIA DE MATRICES

ALBERT, A. A.: *Introduction to Algebraic Theories*. Chicago, The Univ of Chicago Press, 1941.
FRAZAR, R. A., DUNCAN, R. J., y COLLAR, A. R.: *Elementary Matrices and Applications to Dynamics and Differential Equations*. 3.ª edic., Camb The University Press, 1946.
MACDUFFEE, C. C.: *The Theory of Matrices*. 2.ª edic. New York, Cl Publishing Co., 1946.
—*Vectors and Matrices*. Carus Mathematical Monograph N.º 7. Mer Wisconsin, The Mathematical Association of America, 1943.
PERLIS, S.: *Theory of Matrices*. Cambridge, Adisson-Wesley Press, 1952.

TEORIA DE GALOIS

ARTIN, E.: *Galois Theory*, 2.ª edic. Notre Dame Mathematical Lecture, Notre Dame, Indiana, 1944.
STEINITZ, E.: *Algebraische Theorie der Körper*. Con Apéndices, por R. y H. Hasse. Leipzig, W. de Gruyter, 1930; y New York, Chelsea Publi Co., 1951.

ALGEBRAS LINEALES Y ANILLOS

ALBERT, A. A.: *Structure of Algebras*. New York, American Mathematica ciety (Colloquium Publications, Vol. 24), 1939.
ARTIN, E., NESBITT, C. J., y THRALL, R. M.: *Rings with Minimum Cond* Ann Arbor, University of Michigan Press, 1944.
JACOBSON, N.: *The Theory of Rings*. New York, American Mathematica ciety (Mathematical Surveys, N.º II), 1943.
MCCOY, N. H.: *Rings and Ideals*. Carus Mathematical Monograph N.º 8, B lo, N. Y., The Mathematical Association of America, 1948.

TEORÍA DE LA VALORACIÓN

SCHILLING, O. F. G.: *The Theory of Valuations*. New York, American M matical Society (Mathematical Surveys N.º IV), 1950.

GEOMETRÍA ALGEBRAICA

CHEVALLEY, C.: *Introduction to the Theory of Algebraic Functions of One riable*. New York, American Mathematical Society (Mathematical Sur N.º VI), 1951).
HODGE, W. D. D., y PEDOE, D.: *Methods of Algebraic Geometry*. Camb University Press, Vol. I, 1947; Vol. II, 1952.
WALKER, R. J.: *Algebraic Curves*. Princeton, Princeton University Press;

LÓGICA

ROSENBLOOM, P. C.: *The Elements of Mathematical Logic*. New York, I Publications, Inc., 1950.
ROSSER, J. B.: *Logic for Mathematicians*, New York, McGraw-Hill, 1953.
TARSKI, ALFRED: *Introduction to Logic*. New York, Oxford University I 1939.

TEORÍA DE RETÍCULOS

BIRKHOFF, GARRET: *Lattice Theory*. 2.ª edic. New York, American Mathem Society (Colloquium Publications, Vol. 25), 1948.

INDICE ALFABETICO